

Db 88 -----IleGlnGlyArgAspHisPro 94
QY 191 CATAACACGGGCTATCCGACATCTACCTGCTGACGACCCCGAGAGCCCTTGGCCGC 250
Db 95 -----GlyGlnSerGlyPheHisLeuLeuGlyAlaSerGluAspAlaPheValAla 111
QY 251 CGCGCGCCCTTATCGAATTCGCGAACAACACGCTCGATTGCAATCTACTACATTTGGCGC 310
Db 112 ArgAlaAlaLeuIleArgAlaAlaGlnArgSerLeuAspIleGlnTyrTyrIleValHis 131
QY 311 AACGACATTTCCGGCAGGCTCTGTTCAACCTCATGTACCTTGGCGGAGAACGGGGTG 370
Db 132 AspGlyLeuThrArgAlaLeuAlaTyrGluLeuLeuIleValAlaAspArgGlyVal 151
QY 371 CGCGTACGCCCTCTTTGCGACGCAACAACACGCGGGTTGGACGATCTCCTGCTGCC 430
Db 152 ArgValArgIleLeuIleAspThrAlaSerAspGlyTyrPheTyrGluIleGlyVal 171
QY 431 CTCGACAGCCATCCCAATATCGAAGTGCCTGTTCACCCCTTCTGCTTACGCAAA--- 487
Db 172 LeuSerAlaHisProAsnIleGlnValArgLeuPheAsnProLeuHisLeuGlyArgAla 191
QY 488 -----TGGCGGCACTCGGCTACTGACCGACTTCCCGCCCTCAACCGCGCATG 538
Db 192 ThrGlyIleThrArgGlyValGlyArgLeuPheAsnLeuSerGlnGlnHisArgArgMet 211
QY 539 CACAACAAATCCTTTACCGCGCAACAACCGCGGCTTACCTACTCGCGGACGCAATATCGC 598
Db 212 HisAsnLysLeuTyrPheAlaAspGlyThrAlaIleValGlyGlyArgAsnLeuGly 231
QY 599 GACGAATACTTCAAGTGGTGAGGACACCGTTTTCGCGGACCTGGACATCCTCGCCACC 658
Db 232 AspGluTyrPheAsnAlaLysProGluMetAsnPheThrAspLeuAspLeuLeuGlyVal 251
QY 659 GGCAGCGTCTCGCGGAAGTATCCGACGACTTCCACCGCTACTGGGCAAGCCATTCGCC 718
Db 252 GlyProIleAlaAsnGlnLeuSerHisPheAspGlnTyrTrpAsnSerAlaIleSer 271
QY 719 CACAACGCCACG-----CGCATCATCCGCGGCGCAACATCGCGAAGGCTT 766
Db 272 ArgProIleGluAspPheLeuTyrPheArgAlaProTyrProGlyGluLeu----- 287
QY 767 CAAGCATCTGGATACAACGAGCAACATCCAGACCGCTCTGCTGCGCTAC-----CGC 820
Db 288 -----GluSerAlaArgLysLeuGlnArgTyrLeuArgLys 300
QY 821 GAAACCTCGAACAGTCG-----CCCTCTACCAAAATAACAGACGGGACGC 868
Db 301 GluSerValLysGluSerGlyTyrIleArgHisLeuPheAspArgGlyAspGlnProArg 320
QY 869 ATC---GACTGGCAGAGC-----GTCCAAACCCGCTGTATCAGCAGCACCCCT 913
Db 321 LeuGlyAsnTyrPheGluAsnLeuThrTrpAlaArgAlaGluAlaIleTyrPheAlaPro 340
QY 914 GCAAAAGGATCGACCGGACCGCGCCGCAAAACCGCG-----ATTGCGGGAGGCTGCAA 967
Db 341 LeuLysValLeuSerArgGlyGluProAspProHisLeuLeuSerProHisLeuAla 360
QY 968 GACGGCTCAACAGCCGGAAGAGCTATCTGTTTCCACTTATTCGCTCCCTACA 1027
Db 361 GlyLeuPheLysGlyValGlnLysGluLeuIleValSerAlaTyrPheValProAla 380
QY 1028 AAATCCGGCACAGACGACTGGCAAAACTGGTGAGGACGCGCATAGAGTTACCGCTCG 1087
Db 381 LysAspGlyLeuAsnTyrLeuThrGlyLysAlaAspSerGlyValArgValArgLeuLeu 400
QY 1088 ACCAACTCGGTACAGCGACGAGTTCGCGCGCTCCATTCGCGTACGTCACCAATACCGA 1147
Db 401 ThrAsnSerLeuGluAlaThrAspValProAlaValHisAlaGlyTyrAlaProTyrArg 420
QY 1148 AAACCGTGTCAAGCGCGCATCAAACTCTAGAGCTGCAACCAACCATCGCGTCC 1207
Db 421 MetAlaLeuLeuGluHisGlyValLysLeuTyrGluLeuArgAlaAsnProAspGlnPro 440

QY 1208 GCCACA---AAGACAAAGCGCTGACCGGAGCTCCGTAACACGCTGCTGCAAAAACC 1264
Db 441 LeuSerGlyAlaProTyrArgLeuHisGlySerSerSerAlaSerLeuHisSerLysAla 460
QY 1265 TTCATTGTGGACGGCAAAACGATCTTCATCGGCTCATTTCAACCTCGACCCCGTTCCGCA 1324
Db 461 MetValPheAspArgArgLysValPheIleGlySerPheAsnPheAspProArgSerIle 480
QY 1325 CGGCTCAATACCGAATGGCGTCTCATCGAAGCCCAAAATCGCAGAACACATGGAG 1384
Db 481 LeuTrpAsnThrGluValGlyValIleValAspSerProLeuLeuAlaGluGlnValArg 500
QY 1385 CGC---ACCCTCGCGCATACACACCGAATAGCCCTACCGCTTACCTCGCAAAACAC 1441
Db 501 GlnLeuAlaLeuGluGlyMetAlaProSerValSerTyrGlnValArgIleAspArgSer 520
QY 1442 AAC-----CGCCTGCAATGGCAGAT-----CCGCGCACCCGAAACCTACCCG 1486
Db 521 GlySerArgProLysLeuValTyrIleAspGluArgAspGlyArgProGlnValLeuArg 540
QY 1487 AAGAACCCGAAAGCAACTTTGAAACGATCGCGCAAAATCTATCCTGCTGCTGCC 1546
Db 541 HisGluPro---GlySerLeuTyrArgLeuAsnAlaTyrValAlaGlyMetIleGly 559
QY 1547 ATCGAAGGTTTATTA 1561
Db 560 LeuGluLysMetLeu 564

RESULT 3
US-09-328-352-5559
; Sequence 5559, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 5559
; LENGTH: 271
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-09-328-352-5559

Alignment Scores:
Pred. No.: 3,59e-31 Length: 271
Score: 493.50 Matches: 102
Percent Similarity: 63.9% Conservative: 47
Best Local Similarity: 43.8% Mismatches: 73
Query Match: 17.3% Indels: 11
DB: 2 Gaps: 4

US-10-665-990A-13 (1-1561) x US-09-328-352-5559 (1-271)

QY 59 TCCTTTTATGCTCTCTCTGTTTCATGTTCTTATGTTGCTCC----- 103
Db 37 SerSerPheAlaValAlaLeuThrGlyCysSerThr---LeuProLysHisSerProGlu 55
QY 104 CCATCTGGAAGACGGAGGAAAGCGCTCATTTCAATACTTCCAAACCTGCTCTCTCGGAC 163
Db 56 ProIleGln-----TyrAlaArgAspIleAspThrSerGlnThrSerLeuSerLys 72
QY 164 AACATCTCTCAATCCGCGACACACCCCTCATACACACGGGCTATCCGACATCTACCTGCTC 223
Db 73 IleIleThrProLeuArgGlu-----LysAsnProAsnLeuThrGlyTyrHisLeuLeu 90
QY 224 GAGCACCCCAAGAACCCCTTGGCGCGCGCGCTTATCGAATCTCGCGAACACAGC 283
Db 91 AsnAspProLeuGluAlaLeuAlaAlaArgLeuArgLeuIleAspLysAlaGluLysThr 110

QY 284 CTCGATTTCATACATCTTGGCGCACGACATTTCCGCGAGCGTCTGTTCACCTC 343
|||||
Db 111 LeuAspLeuGlnTyrTyrTrpAspAsnAspLysValGluAlaLeuAlaLeuHisAla 130
|||||
QY 344 ATGTACCTTGGCGGAGCGCGCTGCGCTACGCTCTGTGGACGACAAACACAGC 403
|||||
Db 131 IleIleArgAlaAlaAspArgGlyValLysValArgLeuLeuIleAspAsnAsnAla 150
|||||
QY 404 CGCGGTTGGACGATCTCTCTCGCTCGCCCTCGACGCCATCCCAATATCGAAGTGGCGCTG 463
|||||
Db 151 LysLysMetGluGlyValLeuLeuAlaLeuSerGlnHisLysAsnIleGluValLysLeu 170
|||||
QY 464 TTCACCCCTTCTGCTACGCAATGGCGCGCACTCGGCTACTCGACGACTTCCCGCGC 523
|||||
Db 171 PheAsnProTyrArgPheArgLysTyrArgAlaMetAspMetIleLeuAspLeuLysArg 190
|||||
QY 524 CTCACCGCGCGATGCACAAACAACTTTTACCGCGCAACCGCGCCACCATCTCGGC 583
|||||
Db 191 IleAsnArgArgMetHisAsnLysSerPheIleAlaAspAsnGlnValAlaLeuIleGly 210
|||||
QY 584 GGACGCAATATCGCGACGAATCTTCAAAGTCGTGAGGACACCGCTTTTCGCCGACCTG 643
|||||
Db 211 GlyArgAsnMetThrAsnGlnTyrTyrAsnValSerAspSerTyrGlnPheSerAspVal 230
|||||
QY 644 GACATCTCGCACCGGAGCGTCTCGCGGAAGTATCGCACGACTTCGACCGCTACTGG 703
|||||
Db 231 AspValMetLeuValGlyAlaAlaValAspAspIleValAsnSerPheValAspTyrTrp 250
|||||
QY 704 GCAAGCATTTCGCCACAAACGCCAGCGCATCATCCGC 742
|||||
Db 251 AsnHisGluTyrAlaTyrSerValGlnSerIleValArg 263
|||||

RESULT 4

US-09-328-352-5591
; Sequence 5591, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 5591
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-09-328-352-5591

Alignment Scores:
Pred. No.: 275 Length: 275
Score: 478.50 Matches: 100
Percent Similarity: 61.9% Conservative: 51
Best Local Similarity: 41.0% Mismatches: 78
Query Match: 16.8% Indels: 15
DB: 2 Gaps: 5

US-10-665-990A-13 (1-1561) x US-09-328-352-5591 (1-275)

QY 866 CGCATCGACTGGCAGAGCGTCCAAACCCGCTGATCAGCGACACCCCTGCAAAAGACTC 925
|||
Db 35 LysPheAspTrp-----ValLysAlaGluValValLysAspSerProAspLysIleArg 52
|||
QY 926 GACCGCGACCGCGCAACCGCGGATTCGCGGAGCGTCAAGACCGCTCAAAACAGCCC 985
|||
Db 53 SerLysAlaLysLysGluGluHisLeuAsnPheGlnLeuIleAsnHisLeuGluLysPro 72
|||
QY 986 GAAAAAGCGTCTATCTGTTTCAACCCCTATTCGTCCTCCCTACAAAATCCGCGCAGACGCA 1045
|||
Db 73 GluSerAsnValAspLeuIleSerAlaTyrPheIleProGluLysGlnGlyAlaLysIle 92
|||
QY 1046 CTGGCAAAACTGGTGCGAGCGCATAGAGTTTACCGTCTGTGACCAACTCGCTACAGGCG 1105
|||

Db 93 LeuSerThrLeuAlaLysGluGlyValGluValArgValLeuThrAsnSerPheLysAla 112
|||
QY 1106 ACCGACGTTGGCGGCTCCATTCCGGCTACGTCAAATACCGAAAAACCCCTGCTCAAAGCC 1165
|||
Db 113 AsnAspValAlaValAlaHisAlaPheTyrGlyLysTyrArgLysGluLeuLysAsn 132
|||
QY 1166 GGCATCAAACTCTACGAGCTGCAACCC-----AACCAT 1198
|||
Db 133 GlyValGlnLeuTyrGluPheLeuProThrProAspLysArgAspLeuAsnLysAsnThr 152
|||
QY 1199 GCGCTCCCGCCACAAAGACAAA---GGCTGACCGCGAGCTCCGTAAACACGACCTGCAT 1255
|||
Db 153 AspGluLeuAlaThrLysAlaLysValAsnMetLysGlyLeuSerArgSerSerLeuHis 172
|||
QY 1256 GCCAAAACCTTCATTGTGGACGGCAACGCATCTTCATCGGCTCATCAACCTCGACCCC 1315
|||
Db 173 ThrLysLeuMetAlaLeuAsp---GluGlnValPheIleGlySerPheAsnPheAspPro 191
|||
QY 1316 GCTTCGCGACGGCTCAATACCGAAATGGCGTCTCATCGGCTCATCAACCTCGACCCC 1375
|||
Db 192 ArgSerAlaTyrLeuAsnThrGluIleGlyValIleLeuAspSerProSerLeuAlaLys 211
|||
QY 1376 CAGATGGAGCGCACCTCGCGATACACACCCGGAATACGCCTACCGGTTACCCCTCGAC 1435
|||
Db 212 ThrIleHisThrMetAspGluAsnLeuAsnLysTyrAlaTyrLysLeuLysLeuAsp 231
|||
QY 1436 AAACACAAACCGCTGCAATGGCAGCAT-----CCGCGCACCCGAAAAACCTTACCCGAAAC 1489
|||
Db 232 ProAsnAsnHisIleTyrTrpGlnGlnThrProLysGlyProValIleTyrLysLys 251
|||
QY 1490 GAACCGGAGCCAACTTTGGAAACGCATCGCCGAAAAATCCTATCCTGCTGCCCATC 1549
|||
Db 252 GluProGluMetLysTrpGlnLysAlaGlyMetLysLeuLeuSerTrpLeuProLeu 271
|||
QY 1550 GAAGGTTTATTA 1561
|||
Db 272 GluGlyPheMet 275
|||

RESULT 5

US-09-328-352-7685
; Sequence 7685, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 7685
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-09-328-352-7685

Alignment Scores:
Pred. No.: 321 Length: 321
Score: 409.00 Matches: 100
Percent Similarity: 50.0% Conservative: 44
Best Local Similarity: 34.7% Mismatches: 110
Query Match: 14.3% Indels: 34
DB: 2 Gaps: 6

US-10-665-990A-13 (1-1561) x US-09-328-352-7685 (1-321)

QY 74 CTTCCTGTTTCATGTTCTTTCATGGTTG-----CCCCCACTGGAAGACGCGGAAAGC 127
|||
Db 44 IleValIleLysThrAsnAsnTrpLeuAsnAspGluAlaGluGluLysThrGln--- 62
|||
QY 128 CGTCATTTTCAATACTTCCAAACCTGTCTCTCTGGACAAACATCTCTGCAAAATCCGGCACACC 187
|||

```
Db 62 ----- 62
188 CCTCATAAAGCGGCTATCGACATCTACCTGCTCGACGACCCCGCCAGAGCCCTTGC 247
Db 63 -----GlnGlyLeuThrAlaPheValAlaLeuAspAlaPheMetSerIleAla 79
248 GCCCGCGCGCTTATCGAATCTCGGAAACACACGCTCGATTTGGCAATACATCTTGG 307
Db 80 SerArgIleTyrLeuIleArgAsnAlaLysGluThrIleAspLeuGlnTyrTyrIleTrp 99
308 CGCAACGACATTTCCGCGAGCTGCTGTTCACCTCATGTACCTTGGCGAGAGCGGC 367
Db 100 ThrAsnAspPheValGlyAsnLeuIleLeuHisGluLeuLeuLysAlaAlaAspArgGly 119
368 GTGCGCGTACCTGCTGCTTGGACAGCAACACGCGCGGTGGAGCATCTCCTGTC 427
Db 120 IleLysValArgLeuLeuIleAspAsnGlnGlyLeuLysLeuAspGlyIleLeuArg 139
428 GCCCTCGACAGCCATCCCAATATCGAAGTGGCGCTGTTCACCCCTTCTGCTACGCAAA 487
Db 140 SerLeuLeuGlnHisThrAsnPheGluIleArgLeuPheAsnProTyrLysPheArgTyr 159
488 TGGCGCGCATCTCGGCTACCGCATCTCCCGCGCTCAACCGCGCATGACACAA 547
Db 160 LeuArgIlePheAspTyrLeuPheArgPheLysLysValAsnHisArgMetHisAsnLys 179
548 TCCTTTACCGCGGACAAACCGCGCACCATCTCGCGGAGCAATATCGGCGACGATAC 607
Db 180 LeuIleAlaAspAlaSerIleAlaValThrGlyGlyArgAsnIleSerSerGluTyr 199
608 TTCAAAGTCGGTAGGACACCGTTTTCGCGACCTGGACATCTCTCGCCACCGGAGCGTC 667
Db 200 PheGluAlaSerSerLysPheGlnPheThrAspMetAspIleLeuPheTyrGlyHisAla 219
668 GTCCGCGAAGTATCGCAGGCTTCGACGCTACTCGGAGAGCCATTCGCGCCACAGCGC 727
Db 220 ValArgHisAlaGlnAlaValPheThrAspPheTrpGluSerThrLeuSerValAsnAla 239
728 ACGGCGCATCTCCGAGCAACATCGGCAAG--GGTCTTCAAGCACTCGGATACAAC 784
Db 240 ThrGluIlelle-----GlyThrCysAlaGluHisIleLysAlaLeuArgGluHis 257
785 GACGAACA-----TCCAGACGCGCTCTCGCTACCGCGAAACCGTCGACACAG 835
Db 258 TyrGluGlnLeuHisHisGluAspHisSerLeuThrGluAspLysLeuTyrAspAlaGln 277
836 TCGCCCTCTACCAAAAATACAGCGGAGCATCGACTGCGAGCGCTCCAAACCCGC 895
Db 278 SerTyrLeuLysGluLeuLeuGluHisAsnProIleGlnTrpSer-----LysAlaHis 295
896 CTGATCAGCGACACCCCTGCAAAA 919
Db 296 PheValAlaAspSerProLysLys 303
```

```
RESULT 6
US-09-538-092-1258
; Sequence 1258, Application US/09538092
; Patent No. 6753314
; GENERAL INFORMATION:
; APPLICANT: Giot, Loic
; TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same
; FILE REFERENCE: 15966-542
; CURRENT APPLICATION NUMBER: US/09/538,092
; CURRENT FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: 60/127,352
; PRIOR FILING DATE: 1999-04-01
; PRIOR APPLICATION NUMBER: 60/178,965
; PRIOR FILING DATE: 2000-02-01
; NUMBER OF SEQ ID NOS: 1387
; SOFTWARE: CurapadSeqFormatter Version 0.9
; SEQ ID NO 1258
; LENGTH: 5179
```

```
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (0)...(0)
; OTHER INFORMATION: Polypeptide Accession Number Q02817
US-09-538-092-1258

Alignment Scores:
Pred. No.: 3,86e-18 Length: 5179
Score: 336.50 Matches: 141
Percent Similarity: 33.6% Conservative: 55
Best Local Similarity: 24.1% Mismatches: 198
Query Match: 11.8% Indels: 190
DB: 2 Gaps: 22

US-10-665-990A-13 (1-1561) x US-09-538-092-1258 (1-5179)
QY 119 ACGGAAAGCGTCAATTTCAAT-----ACTTCCAAACCTGTCTCTCGACAAC 166
Db 1248 ThrValGluLysHisPheAsnIleCysSerIleThrThrArgPro-SerThrLeuThrTh 1267
QY 167 ATCTCTCAAAATCCGGCACACCCCTCATPAAACAACGGGCTATCCGACATCTACCTCTCGAC 226
Db 1267 rPheThrThrIleThrLeuProThrThrProThrSerPheThrThrThrThrThrTh 1287
QY 227 GACCCCAACGAGCCCTTGGCGCAACGACATTTTCGGCAGGCTGTGTTCAACCTCA-- 286
Db 1287 rThrProThr-----SerSerTh 1293
QY 287 GATTTCGAATACATCTTGGCGCAACGACATTTTCGGCAGGCTGTGTTCAACCTCA-- 344
Db 1293 rValLeuSerThrTh-----ProLysLeuCysCysLeuTrpSerAs 1307
QY 344 ----- 344
Db 1307 pTrpIleAsnGluAspHisProSerSerGlySerAspAspGlyAspArgGluProPheAs 1327
QY 345 -----TGTAACCTTGGCGCAGACCGGGCGGTGGCGTACGCC-- 380
Db 1327 pGlyValCysGlyAlaProGluAspIleGluCysArgSerValLysAspProHisLeuSe 1347
QY 380 ----- 380
Db 1347 rLeuGluGlnHisGlyGlnLysValGlnCysAspValSerValGlyPheIleCysLysAs 1367
QY 380 ----- 380
Db 1367 nGluAspGlnPheGlyAsnGlyProPheGlyLeuCysTyrAspTyrLysIleArgValAs 1387
QY 381 -TCGCTGTTGG-----ACGACAACAAC 400
Db 1387 nCysCysTrpProMetAspLysCysIleThrThrProSerProThrThrThrProSe 1407
QY 401 ACGCGCGGTGGACGATCTCTGCTGCGCCCTCGACAGCCATCCCAATATCATGAGTGC 460
Db 1407 rProProProThrThrThrThrLeuProProThrThrThrPro----- 1422
QY 461 CTGTTCAACCCCTTCTGCTACGCAAAATGGCGCGCATCGGCTACCTGACCGACTTCCTCC 520
Db 1423 ----SerProProThrThrThrThrThrThrThrThrThrThrThrThrProPr 1441
QY 521 CGCCTCAACCGCGCATGACACAATACTCTTTACCGCGCAACCGCGCCACCATCTC 580
Db 1441 oIleThrThr-----ThrThrThrProLeuProThrThrThrProSerPro----- 1456
QY 581 GCGCGAGCGAATATCGCGACGAATACTTCAAAAGTCGGTGAGGACACCGTTTCGCGCGAC 640
Db 1457 -----ProIleSerThrThrThrThrProProProThrThrThrProSerProTh 1474
QY 641 CTGACATCTCCGACCGCGAGCGTGTGGCGAAGTATCGACGACTTCGACCC----- 695
Db 1474 rThrThrProSerProProThrThrThrProSerProProThrThrThrThrThrPr 1494
```

```
QY 696 -----OCTACTGGGCAAGCCATTCCGCCCAACAGCCGCGCATCATCGCAGCGGC 748
Db 1494 oProProThrThrThrProProProMetThrThrProleThrThrProAlaSerTh 1514
QY 749 AACATCGGCAAGGCTTCAGAGCACTCGATACACGACGAAACATCCAGACAGCGCTC 808
Db 1514 rThrThrLeu-----ProProThrThrThrThrProSerProProThrThrTh 1529
QY 809 CTGCGCTACCGGCAACCGTGAAC-----AGTCGCCCTCTACCAAAAAATACAGACG 862
Db 1529 rThrThrThrProProProThrThrThrProSerProProThrThr----- 1544
QY 863 GGACGATCGACTGGCAGAGCGTCAAAACCCGCTGATCAGCAGCACCCCTGCAAAAGGA 922
Db 1545 -----ThrProleThrProPr 1550
QY 923 CTCGACCGGACCGCGCAACCGCGGATTCGGGGAGGCTGCAAGACGCGCTCAAAACAG 982
Db 1550 oThrSerThrThrThr-----LeuPro-----ProThrThrThrProSe 1563
QY 983 CCGCAAAAAAGCTATCTGTGTTTCAACCCTATTTCTGTCCTTACAAAATCCGGCACAGAC 1042
Db 1563 rProProProThrThrThrThrProProProThrThrThrProSerProProThrTh 1583
QY 1043 GCATGGCAAAACTGGTCAGGACGCGCATAGAGTTACCGTCTGACCA-----AC 1093
Db 1583 r-----ThrThrProSerProProThrIleThrTh 1593
QY 1094 TCGCTACAGCGACGAGTTGCCCGCTCCATTCGGCTAGCTCAAT----- 1142
Db 1593 rThrThrProProProThrThrThrProSerProProThrThrThrThrThrProPr 1613
QY 1143 -----ACGAAAACCGCTGCTCAAAAGCGGATCAAACTC 1177
Db 1613 oProThrThrThrProSerProProThrThrThrProleThrThrProThrSerThrTh 1633
QY 1178 TAGAGCTGCAACCCCAACATGCGTCCCGCCCAAAAAGACAAAGAGGCTGACCGGACG 1237
Db 1633 rThrLeuProProThrThrThrProSerProProThrThr-----ThrThrTh 1650
QY 1238 TCGGTAAACGAGCTGCATGCGCAAAACCTTCATTGTGGAGCGCAACGATCTTCATCGC 1297
Db 1650 rProProProThrThrThrProSerProProThr---ThrThrThrProSerProPr 1669
QY 1298 TCATTCAACCTCGACCCCGTTCCGCGCTCAATACGAAATGGCGTGTCTATCGAA 1357
Db 1669 eThrThrThrThrThrProProProThrThrThrProSerSerProleThrThrThPr 1689
QY 1358 AGCCCCAAATCGCAGAACAGATGGAGCGCACCTCGCGGATACACACCCGAATACGCC 1417
Db 1689 oSerProProThr---ThrThrMetThrThrThrProSerProThrThrThrProSerSerPr 1708
QY 1418 TACCGGTTACCTCGACAAACACACGCGCTGCAATGCGAGATCCCGCCACCGAAAA 1477
Db 1708 oile-----ThrThrThrThrThrProSerSerThrThrThrProSerProPr 1725
QY 1478 ACCTACCGGAAGAACCCGAACTTTGGAAACGATCGCGCAAAATCCTATCC 1537
Db 1725 oThrThrMetThrThrProSerProThrThrThrProSerProThrThrThrMetTh 1745
QY 1538 CTGCTGCCCA 1547
Db 1745 rThrLeuPro 1748
```

RESULT 7

```
US-09-252-991A-21920
; Sequence 21920, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
```

```
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252.991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 21920
; LENGTH: 822
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-21920

Alignment Scores:
Pred. No.: 1,19e-17 Length: 822
Score: 326.50 Matches: 188
Percent Similarity: 32.5% Conservative: 52
Best Local Similarity: 25.5% Mismatches: 195
Query Match: 11.4% Indels: 303
DB: 2 Gaps: 40
```

```
US-10-665-990A-13 (1-1561) x US-09-252-991A-21920 (1-822)

QY 103 CCCACTGGAAGACGGAGCGGAAGCCGTCATTTCAATACTTCCAAACCTGTCT- 156
Db 86 ProAlaGlyArgProAspLeuGlnProGlyArg- 100
QY 157 -----CCTGGACAACATCTCGCAATCCGGCACACCCCTCATACAACGCGCT 204
Db 101 AlaGlyArgGlnProGlyGlnProProGlyHisAlaProGlyProGlyArgGlnProAla 120
QY 205 ATCCGACATCTACCTGCTCGACGACCCGCCAGAGCCCTTCGCCGCCGCCGCTTAT 264
Db 121 AlaTyrArgTyrAlaGlyGlnProGlyArgProProGlyGlnProGlyArgGlnPro 140
QY 265 CGAATCTGCCAAACACAGCCTCGATTTCATATACTACATTTTGGCGCAACAGCATTTCCGG 324
Db 141 ArgProAlaGluArgGlnProArg-----GlnArgArgArgArg 153
QY 325 CAGGCT-----GCTGTTCAACCTCATGTACCTTGCCTGCGCAGAACGCGG 366
Db 154 ArgAlaGlnArgGlnGlyLeuAlaAspAlaGlyHisArgAlaValArgGlnGlnArg 173
QY 367 CGTGGCGTACGCTGCTGTT----- 387
Db 174 ArgArgHisProGlyAlaValAlaGlyAspSerArgArgAlaArgAlaGlnProAla 193
QY 388 -----GGACGACAAACAA-----CACGCGCGG----- 408
Db 194 GlyProSerLeuGlyAlaGlyArgGlnProHisArgHisArgArgLeuArgGlnPro 213
QY 409 -----GTTGGACGATCTCTGCTCGCCCTCGACGCCATCC----- 444
Db 214 GlyTyrArgProLeuArgGlnArgProAlaGlnProArgProArgProAlaPheProGlnPro 233
QY 445 -----CAATATCGAAGTCCGCT--- 462
Db 234 GlyArgGlyGlySerGlyArgGlnGlyArgArgArgAlaHisArgLeuGlnProGly 253
QY 463 -----GTTCAACCCCTTCGCTACGCAAAATGGCGCAGCTCGGCTACCTACCGCA 513
Db 254 ArgArgAlaGlyGlnProLeuArgProValGlyLysArgLysArgAlaAlaProAlaArg 273
QY 514 -----CTTCCCGCGCTCAACCGCCGCGCATGCA 540
Db 274 ArgArgAspArgGlnGlnArgGlnProAlaArgProArgProGlnArgGlnHisAla 293
QY 541 -----CAACAAATCCTTACCGC-----CGACAA-----CCG 567
Db 294 AlaGlyArgTyrArgProGlnArgLeuArgArgAlaGlyLysArgGlnProGlyPro 313
QY 568 CGCCACCATCTCTCGGCGGACGCAATATCGGCGACCAATA-----CTT 609
```


Db	644	SeThrThrThrGluSerSerSerAlaProValProThrProSerSerSerThr-----Thr	661
Qy	1091	AACTCGCTCAGCGACCGACGAGTTCGCCGCG-----TCCAT	1126
Db	662	GluSerSerSerAla-ProValProThrProSerSerSerThrThrGluSerSerAla	681
Qy	1127	TCCGGCTACGCTAAATACCGAAACCGCTGCTCAAAGCCGGCATCAAACTCTACG	1181
Db	681	aProValThrSerSerThrThrGluSerSerSerAlaProValThrSerSerThrThrG1	701
Qy	1182	-ACTGCAACCCCAACCATGCCGTCGCCGCACAAAACAGACAAAGCCGTGACCGCAGCTCC	1240
Db	701	uSerSerSerAlaProValProThrProSerSerSerThrThrGluSerSerSerAlaPr	721
Qy	1241	GTAACCAAGCTGCATGCGCAAAACCTTCATTGTGGACGGCAACGCATCTTCATCGGCTCA	1300
Db	721	o-----ValProThrProSer-----SerSerThrThrG1	731
Qy	1301	TTCAACGCTCGACCCCGCTTCCGACGCTCAATACCGAAATGGCGTGTCTATCGAAAGC	1360
Db	731	uSerSerSerAlaProValProThrProSerSerSerThrThrGluSerSerSerAlaPr	751
Qy	1361	CCCAAAATCCGACAAACAGATGGAGCGCACCTCGCGATACCAACCCGGAATACG-----	1415
Db	751	oValThrSerSerThrThrGluSerSerSerAlaProValProThrProSerSerSerTh	771
Qy	1416	-----CCTACCGGTTTACCTCGACAAACAAACAC---GCCTGCAA	1453
Db	771	rThrGluSerSerSerAlaProValProThrProSerSerSerSerThrThrGluSerSerSe	791
Qy	1454	TGGCAGCATCCCGCCACCGAAAAAACCCTACCGAAACGAAC-----CCGAA	1498
Db	791	raLaProValProThrProSerSerSerThrThrGluSerSerValAlaProValProTh	811
Qy	1499	GCCA-----AACTTGGAAACGCATCGCCGCAAAATCC	1532
Db	811	rProSerSerSerSerAsnileThrSerSerAlaProSerSerThr	826
RESULT 9			
US-09-252-991A-17112			
; Sequence 17112, Application US/09252991A			
; Patent No. 6551795			
; GENERAL INFORMATION:			
; APPLICANT: Marc J. Rubenfield et al.			
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUD			
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS			
; FILE REFERENCE: 107196.136			
; CURRENT APPLICATION NUMBER: US/09/252,991A			
; CURRENT FILING DATE: 1999-02-18			
; PRIOR APPLICATION NUMBER: US 60/074,788			
; PRIOR FILING DATE: 1998-02-18			
; PRIOR APPLICATION NUMBER: US 60/094,190			
; PRIOR FILING DATE: 1998-07-27			
; NUMBER OF SEQ ID NOS: 33142			
; SEQ ID NO 17112			
; LENGTH: 655			
; TYPE: PRT			
; ORGANISM: Pseudomonas aeruginosa			
US-09-252-991A-17112			
Alignment Scores:			
Pred. No.: 8,23e-16 Length: 655			
Score: 303.00 Matches: 155			
Percent Similarity: 33.5% Conservative: 37			
Best Local Similarity: 27.1% Mismatches: 179			
Query Match: 10.6% Indels: 202			
DB: 2 Gaps: 29			
US-10-665-990A-13 (1-1561) x US-09-252-991A-17112 (1-655)			
Qy	157	CCTGGACAAACATCTTGCAAAT-----CGGCACACCCCTCATATAACAA	198
Db	140	ProAlaGluGluProAlaGlyValArgGluLeuProArgHisProAlaGlyGly	159

199	QY	CGG-----GCTATCGGACATCTACTGCTCGACGACCCCGACGAGCCCTTGC	246
160	Db	ArgAspProAlaGlyArgLeuArgLeuValProArgArgProArgHis	177
247	QY	CGCCCGCCGCGCCCTTATCGGAATCTGCCGACACACAGCCTCGATTGTCATCTACATTG	306
178	Db	-----TrpA ^g GlyLeuGlyArgValArgAlaValProArgHisAspLeuGlyArgAla	195
307	QY	CGGCAACGACATTTCCGCGCAGGTGCTTCAACCTCATGTACTTCCGCGAGAACGGG	366
196	Db	AspGluAlaArgLeuArgArgCysProAlaProGlyLeuAlaProArgArgThrLeu	215
367	QY	CGTGGCGGTACGCTGCTGTGTGGAGCAACACACGCGCGGTGTGGACGATCTCTGCT	426
216	Db	ArgArgArg-----ProLeuHisArgProGlnGlyAlaArgProGlyArgSerAlaGly	233
427	QY	CGCCCTCGACAGCCATCCCAATATCGA-----AGTGCCTGCTTTCAACCC	471
234	Db	AlaAlaGlyArgGlyLysAlaGlnArgLeuHisLeuGluProValAlaArgValArgArg	253
472	QY	CTTCGTCTAGCGAAATGGCGGCACTCGGTACTCTGAC-----CGATTTCCCGCGCT	525
254	Db	LeuArgAlaValArgGlnValArgGlyAspValProGlyValAlaArgArgProAlaAla	273
526	QY	CAACCGCGCATGACACAATACTTTACCGCGCACAAACCGCGCCACCATCTACTCGCGG	585
274	Db	GluProGluGluAla-----AspProGlyHis-----	282
586	QY	ACGCAATATATCGCGCAGCAATACTTCAAAATCGGTGAGGACACCGTTTTTCGCGCACCTGGA	645
283	Db	-----GlyHis-----ArgProGly	287
646	QY	CATCTCGCCACCGGACGCTGCTGGGAAAGTATCGCACGACTTCGACCGCTACTGGCG	705
288	Db	ArgArgGlnArgGluValArgArgGlnProLeuSer-----ArgGlnThrAlaGly	305
706	QY	AAGCCATTCGCG-----CCACAACGC	726
306	Db	ArgAlaArgArgArgProAlaProAlaAspArgArgAlaGlyTrpGlnGlyProGlyArg	325
727	QY	CACGCGCATCTCGCAGCGGCAACATCTCGGCAAGGCTCTTCAAGCACTCGGATACAACGA	786
326	Db	ArgArgHisProValValLeuHisLeuProGly-----LeuArg	339
787	QY	GGAAACATCCAGACACCGGCTCTCGGCTACCGGAAACCGTGGACAGTCGCCCTCTA	846
340	Db	ArgGlyValProAspAspArgAlaArgArgArgHisArgArgHis--AlaProPro	358
847	QY	CCAAAAATACAGACGGGCGCATCGACTGCGACAGCGT-----CCAAACCCG	894
359	Db	Pro-----AspProGlyGluArgArgHisAlaGlnGlnGlyArgGlyValGlyGln	376
895	QY	CCTGATCGGACACCCCTCGAAAGGACTCGACCG-----	930
377	Db	ProAspArgHisArgGlnProArgArgPheargProArgArgProValGluLeuGlyGly	396
931	QY	CGACCGCGCAACCCCGGATTGCGGGAGGCTCGAAGCG-----	972
397	Db	ArgProGluProAlaAlaAspGluArgGluGluArgArgArgAlaValLeuGlyArg	416
973	QY	-----GCTCAACAGCCCGAAAAAGCGCTCTA-----	999
417	Db	ArgTrpArgLeuArgHisAlaGlnProAlaHisProAlaArgLeuArgGluAspProGln	436
1000	QY	-----TCTGGTTTCACCCCTATTTCGTCCTCTACAAAATCCGGCACAGACGCTGGC	1050
437	Db	GlySerArgGlyGlyLeuArgArgAlaArgProGlyGlyAlaArgGlnArgArgArgGly	456
1051	QY	AAAACCTGGTCGAGCAGCG-----	1068
457	Db	-----AlaProAlaGlyArgArgGlyTyrLeuProGluProGlyGlnAlaGlnTyrArgAsp	475

QY 1069 -----CATAGACGTTTACCGTCTCGTACCAACTCGCTACAGGC 1104
 Db 476 ProGlyGluIleProLeuHisProHisArgGlnLeuArgProAlaGlnLeuProCysAla 495
 QY 1105 GACCGACGTTGCCCGCGCTCATTCGCGTACGTCATAATACCGAAAAACCGTGTCAAGC 1164
 Db 496 GluGluArgIleArgArgPro-----ArgArgArgLeuProGlyThrAlaProGlnHis 513
 QY 1165 CGGCATCAAACTCTACGAGCTGCACCCACCAACCATGCGTCCCGCCCAAAAAAGACAAAGG 1224
 Db 514 LeuHisArg-----ArgThrAlaAlaArg-----ArgHisAlaAlaThrGly 527
 QY 1225 CTTGACCGGACGCTCCGTAACCGCTGCATGCCAAAACCTTCATTGTGACACGCGCAACG 1284
 Db 528 ProGlnGln-----GlyArgGlnArg 534
 QY 1285 CATCTTCATCGGTTCATTCAACCTCGACCCCGTTCCGCGACGGCTCAATACCGAAATGGG 1344
 Db 535 HisLeuProArgProLeuLeuPro----- 542
 QY 1345 CGTCGTATCGAAAGCCCCAAAATCGCAGAACAGATGGAGCGCACCTCGCCGATACCCAC 1404
 Db 543 -----GlyProLeuGlnArgArgValargGlyAlaAlaGlnArgAlaAlaArg 558
 QY 1405 ACCCGA-----ATACGCTTACCG----- 1422
 Db 559 ThrArgHisArgGlyGlnGlyAsnAlaThrLeuArgLeuProLeuAlaLeuLeuArgArg 578
 QY 1423 -----CGTTACCTTCGACAAACACACACCGCTGCAATGGCAGCA 1461
 Db 579 ArgTrpArgArgAlaAspHisArgHisProArgGlnAlaThrAspSerArgHisAlaHis 598
 QY 1462 TCCCGCCACCGCAAAACCTTACCGAACGACCGCAACCTTTGGAAACGCATCGC 1521
 Db 599 GlyArgHisProArgAsp-----Arg 605
 QY 1522 CGCAAA-----AATCCTATCCCTGCTGCCCATCGA 1551
 Db 606 ArgArgAlaGlyGlyArgArgLeuProAlaValHisArg 618
 RESULT 10
 US-09-252-991A-23696
 ; Sequence 23696, Application US/09252991A
 ; Patent No. 6551795
 ; GENERAL INFORMATION:
 ; APPLICANT: Marc J. Rubenfield et al.
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
 ; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
 ; FILE REFERENCE: 107196.136
 ; CURRENT APPLICATION NUMBER: US/09/252,991A
 ; CURRENT FILING DATE: 1999-02-18
 ; PRIOR APPLICATION NUMBER: US 60/074,788
 ; PRIOR FILING DATE: 1998-02-18
 ; PRIOR APPLICATION NUMBER: US 60/094,190
 ; PRIOR FILING DATE: 1998-07-27
 ; NUMBER OF SEQ ID NOS: 33142
 ; SEQ ID NO 23696
 ; LENGTH: 618
 ; TYPE: PRT
 ; ORGANISM: Pseudomonas aeruginosa
 ; US-09-252-991A-23696

Alignment Scores:		
Pred. No.:	9.67e-16	Length:
Score:	302.00	Matches:
Percent Similarity:	34.8%	Conservative:
Best Local Similarity:	29.0%	Mismatches:
Query Match:	10.6%	Indels:
DB:	2	Gaps:
		38
		618

US-10-665-990A-13 (1-1561) x US-09-252-991A-23696 (1-618)

QY 169 CCTGCAATCCGGCACACCCCTCATAACAACGGGCTATCCGACATCTACCTGCTCGACGA 228


```
Db 4 ProAspHisProArgHisProArg---ArgArgGlnLeuAlaGluLeuArgIleArgArg 22
QY 229 -----CCCCCAGGAAGCCTTGGCCGCCCGCCGCCCTTATCGAATCTGCCGAACA 279
Db 23 ArgThrAlaProGlySerAspArgGlnGlnProArgArgThrProArgIleArgArg 42
QY 280 CAGCCTCGATTGGCAATACTACATTGGCGCAACGACAT----- 318
Db 43 -----HisGlnGlyGlnSerHisArgProAlaHisGlnGlyArg 55
QY 319 -----TTCGCGCAGGCTGCTTCAACCTCAT 345
Db 56 GlnArgGlnProGlyAlaProAlaAlaMetGlyLeuArgArgAla-----TrpProAla 73
QY 346 GTACCTTGGCGAGAACGGCGGTGCGTACGCTGCTGTTGGACGACAAACACGCG 405
Db 74 AlaProCysGlyArgArgArgProAspThr-----GlnLeuArgLeu 88
QY 406 CGGGTTGGACGATCTCTGCTCGCCCTCGACGCCATCCAAATATCGAAGTGGCGCTGT 465
Db 89 ArgProGlnArgGlnSerGlyArgAsnGlnPro-----AlaProVal 103
QY 466 CAACCCCTT-----CGTCTACGCAAAATGCGCGC 495
Db 104 ArgProGlnProGlyLeuArgCysProArgProAlaGlyArgProGluArgSerProArg 123
QY 496 ACTCGCTACCTGACCGACTTCCCGCCTCAACCGCCGATGCAACAA----- 546
Db 124 ArgGlnAspProThrArgLeuArgArgProGlyGlnProHisArgGlyGlnGlyProAla 143
QY 547 -----ATCCTTTACCGCCGACAAACCG----- 567
Db 144 ArgCysHisHisProLeuArgIleArgArgProArgGlnProAspProThrGlyGlnPro 163
QY 568 -----CGCCACCATCTCGCGGAGCAATATCGCGCAGCAATACTTCAAGTCGGTGA 621
Db 164 GlyGlnArgHisHisLeuArgAlaArgArgArg-----GlnArgHisPro 180
QY 622 GCACACCGTTTTCGCGCTCGACATCTCTCCACCGCGCAGCGTCTGCGCGAAGTATC 681
Db 181 ProHisArgArgSerArgCysGlyHis---ArgValSerLeuArgArgProGlnSerAla 199
QY 682 GCACGATTCGACCGCTACTGGGCAAGCATTCCGCCGCAACGCGCAGCGCATCATCGG 741
Db 200 GlyArg---AlaProLeuAlaGluArgPro---GluProArgArgThrVal-----Pro 215
QY 742 CAGCGCAACATCGCAAGGCTTTCACGCACTCGGATACAA---CGACGAAACATCCAG 798
Db 216 LeuArgProHisArgArgArgGlnGlnGlyHisArgProProGlyArgHisArgTrpArg 235
QY 799 ACACGCGCTCTCGCTACCGGAAACCGTCCGACAGTCCGCCCTCTACCAAAAAATACA 858
Db 236 SerArgGlnProGlyValProLeuArgArg----- 245
QY 859 GACGGAGCATCGACTGGCAGAGCGTCCAAACCCCGCTGATCAGGACACCCCTGCAAA 918
Db 246 -----AlaArgGlnProGlyArgAlaGlyAlaGlnHisPro----- 257
QY 919 AGGACTCGACCGCGCGCAACCGCGATTCGCGGGAG---GCTCGAAGACGCGCT 975
Db 258 -----ProArgProAlaAspProAlaArgProArgAspLeuProLeuArgArgGly 274
QY 976 CAAACACCGCGAAAAAGGCTCTATCTGTTTCCACCTATTTCGTCCTTACAAAATCCGG 1035
Db 275 GlnProThrAlaGlyAspArgLeu-----ProLeuArgProArg 287
QY 1036 CAC-----AGACGCACTGGCAAA-----ACTGGTCAGGACCG 1068
Db 288 HisArgLeuProAlaGlnArgArgArgProGlyArgGlnArgAspProGlySerGlyArg 307
QY 1069 CATAGACGTACCGTCTCT-----GACCAACTCGCTACA----- 1101
```

```
Db 308 GlnGlyAlaGluHisProGlyArgAlaAspArgLeuProAlaLeuArgProAlaAlaAla 327
QY 1102 -----GGCGACCGACGTTGCCCGCGTCCATTCCCGCTACGTCAA---ATACCGAAA 1149
Db 328 ProAspLeuGlyGlnArgHisHisSerGlnPro---ArgValArgProGlyLeuProThr 346
QY 1150 ACCGCTGCTCAAGCCCG-----CATCAAACTCTACGAGCTGCACCAACCAA----- 1194
Db 347 AlaAlaAlaGluGlyArgProLeuAlaGluArgLeuProAlaArgCysGlnTrpGlnTyr 366
QY 1195 -----CCATGCGCTCCCGCCCAACAA 1215
Db 367 ProAlaAlaProProGlnProLeuGlyHisProGlyLeuProValArgProAlaGlyPro 386
QY 1216 AGACAAAGCGCTGACCGCAGCTCCGTAAACGAGCTGCATGCCAA---AACTTCATTGT 1272
Db 387 -----ProAspArgGlyThrArgArgProGlyArgAlaGlnLeuCysLeuArg 403
QY 1273 GGACGGCAACCGCAT-----CTTCATCGGCTCATTCAA 1305
Db 404 GlyArgGlnProHisProThrGlnArgGlnProArgLeuArgArgHisArgGln---Gln 422
QY 1306 CCTGACCCCGTTCGCGACGCTCAATAC-----CGAAATGCGGCTCGTCATCGA 1356
Db 423 ProGlyLeuProValArgAlaArgGlnGlnProIleAspArgHisArgAlaSerGly 442
QY 1357 AAGCCCCAAATTCGAGAACAGATGAGCGCACCTCCCGCATACCAACCGCAATACGC 1416
Db 443 AspGlnArgArgArgGlnProHisProGlyProArgArgAlaGlnThrGlyLeuArg 462
QY 1417 CTACCGCTTACCCTCGACAAACACACCGCTCGAATGGCA----- 1458
Db 463 ArgProGlyProProAlaLysArgGlnProArgProArgProAlaGlyArgArgIleProLeu 482
QY 1459 -----CGATCCCGCCACCCGAAA-----AACCTACCC--- 1485
Db 483 GlnArgProArgProAlaHisArgGlnAlaHisProArgGluArgHisHisLeuProLeu 502
QY 1486 -----GAACGAAACCGAAGCCAACTTTGGAAACGCAT 1518
Db 503 TrpSerArgArgProValAlaArgArgSerArgThrArgArgGlnArgProGluThrAla 522
QY 1519 CGCCGCAAAATCTCTATC-----CCTGCTGCC-----CATCGAAGG 1554
Db 523 SerProValLeuProValAlaGlyGlnProAlaAlaGlyHisHisArgArg 539
RESULT 11
US-09-252-991A-28358
; Sequence 28358, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 28358
; LENGTH: 618
; TYPE: PRN
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-28358
Alignment Scores:
Pred. No.: 1,16e-15 Length: 618
Score: 301.00 Matches: 175
Percent Similarity: 34.6% Conservative: 31
Best Local Similarity: 29.4% Mismatches: 196
```

Qy 1039 -----AGACCACCTGGCAAA-----ACTGGTCAGACGCGCAT 1071
Dy 289 ArgLeuProAlaGlnArgArgArgProGlyArgGlnArgAspProGlySerGlyArgGln 308
Qy 1072 AGAGCTTACCGTCCT-----GACCAACTCGCTACA----- 1101
Dy 309 GlyAlaGluHisProGlyTrpAlaAspArgLeuProAlaLeuArgProAlaAlaPro 328
Qy 1102 -----GGCCACCACGCTGTCGCCGCTCCATTCCGCTACGTCAA---ATACCGAAAAACC 1152
Dy 329 AspLeuGlyGlnArgHisHisSerGlnPro---ArgValArgProGlyLeuProThrAla 347
Qy 1153 GCTGCTCAAAAGCCGG-----CATCAAACTCTACGAGCTGCAACCCAA----- 1194
Dy 348 AlaAlaGluGlyArgProLeuAlaGluArgLeuProAlaArgCysGlnTrpGlnTyrPro 367
Qy 1195 -----CCATGCGCTCCCGCCGCCCAAAAAGA 1218
Dy 368 AlaAlaProProGlnProLeuGlyHisProGlyLeuProValArgProAlaGlyPro--- 386
Qy 1219 CAAAGGCGCTACCGCGCAGCTCCGCTAACCGAGCTCGATGCCAA---AACCTTCATTGTGGA 1275
Dy 387 -----ProAspArgGlyThrArgArgProGlyArgAlaGlnLeuArgLeuArgArgGly 404
Qy 1276 CGGCAAAACGCAT-----CTTCATCGGCTCATCTAACACCT 1308
Dy 405 ArgGlnProHisProThrGlnArgGlnProArgLeuArgArgHisArgGln---GlnPro 423
Qy 1309 CGACCCCGCTTCGCGACGGCTCAATAC-----CGAAATGGGCGTCGTATCGAAAG 1359
Dy 424 GlyLeuProValArgAlaArgGlnGlnProIleAspArgHisArgArgAlaSerGlyAsp 443
Qy 1360 CCCCAAAATCGCAGAACAGATGAGCGCACCCTCGCGATACCACACCCGAAATACGCCTA 1419
Dy 444 GlnArgArgArgGlnProHisProGlyProArgArgAlaGlnThrGlyLeuArgArg 463
Qy 1420 CCGCGTTACCTCGACAAACACACCCGCTCGCAATGGCA----- 1458
Dy 464 ProGlyProAlaAlaArgGlnProArgArgProAlaGlyArgArgileProLeuGln 483
Qy 1459 -----CGA 1461
Dy 484 ArgProArgProAlaHisArgGlnAlaHisProArgGluHisHisLeuProLeuArg 503
Qy 1462 TCCCGCCACCGAAAAACCTACC CGAACGACCGAACCGCAAACTTTGGAAACGCGATCGC 1521
Dy 504 SerArgProValAlaArgArgSerArggthrArgArgGlnArgProGluThrAlaSer 523
Qy 1522 CGCAAAATCTCTATC-----CCTGCTGCC-----CATCGAAGG 1554
Dy 524 ProValLeuProValAlaGlyGlnProAlaAlaGlyHisHisArgArg 539
RESULT 12
US-09-252-991A-23598
; Sequence 23598, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 23598
; LENGTH: 507
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-23598

Qy 109 GGAAGAACGGACGGAA

US-10-665-990A-13 (1-1561) x US-09-252-991A-16789 (1-774)

QY	109	GGAGAAGCGACGGAAAGCGCGTCATTTC	CAATACATCTCCAAACCTGCTCTCTCGACACAT	168
DB	58	GLVARGlnasp--LeuProLeuLeuGln--	-----ProProGlySerArg	71

QY	169	CCTGCAAAATCCGCGCACACCCCTCATACAAACGCGGTATTCGACATCTACTCTGCTCGACGCA	228
Db	72	Pro---AspProGlyArgSerArgGlnAlaAlaAspValAlaGluGlyProAlaGlyLys	90
QY	229	CCCCCA-----CGAAGCCCTTCGCGCGCGCGCGCGC	258
Db	91	ProProAlaLeuGlyGlyArgGlnHisArgHisArgArgProGlnGlyProAlaArg	110
QY	259	-----CCTTATCGAATCTCCGAACACAGCCCTCGATTTCGAATA-----	297
Db	111	LeuAlaGlnGlyAlaProPheArg-----ProArgAspProValProPro	125
QY	298	-----CTACATTGGGGGCAACGACATTTCCGGCAGGCTGCT-----	333
Db	126	GlyAlaSerProAspAlaGlyLeuHisArgGlyAlaArgGlyArgProGlyArgHis	145
QY	334	GTTCAACCTCATGCTCTTCGCGCAGAACGGCGGTGCGGTACGCTGCTGTT-----	387
Db	146	AlaArgGlyHisGlyGlnGlyArgArgProAlaGluAspGlnProAlaValAlaGly	165
QY	388	-----GGACGCAACAAACACGCGCGGTGGCA-----CGA	417
Db	166	GlyProGlyHisArgProLeuGlyAspGlyArgGlnValArgGlnAlaGlyLeuArg	185
QY	418	TCTCTGCTCGCCTCGACAGCATCCCAATATCGAAGTGGCTGTTTCAACCCCTTCGT	477
Db	186	ThrGluArgArgAspAlaAlaGlnArgArgThrLeuArgLeuProAlaLeuGly	205
QY	478	CCTACGCAATCGCGCGC-----ACTCGGCTACTGACACGACTTCCCGCGCT	525
Db	206	ProGluArgLeuArgGlnLeuGlnArgGlyThrAlaGlyHisArgHisLeuProGly	225
QY	526	CAACCG-----CCGCATGCA-----CAACAAATCCCT	552
Db	226	GlnProGlyValProGlyProHisGlyLeuAspGlnGlyTrpGlnAspLeuArg	245
QY	553	TAC-----CGCGCAACACCGCGCCACCAT-----ACTCGCGCG	585
Db	246	LeuProArgHisProGlyArgHisArgLeuProHisHisAspGlnArgProArgArg	265
QY	586	ACGCAATATCGCGCAGATCTTCAAGTCGGTGAGGACACCGTTTTCCGCGACCT-----	642
Db	266	AlaArgLeuGlyArgArg-----ArgHisArgGlyGlySerArgHisAlaArgProAla	283
QY	643	-----GGACATCTCTGC-----CACCGGACGGTCTCGCGGAAGTATCGCACGACTT	690
Db	284	GlyIleAspAlaAspProArgGlyHisArgLeuGlnAlaHisArgGlnAlaAlaArgGly	303
QY	691	CGACCGCTACTGCGCAAGCCATTCCGCGCCCAACCGCCACGCGCATCATCGCAGCGGCAA	750
Db	304	His-----HisArgHisArgProGlyAlaAspArgHisAlaAspAlaAlaGln	319
QY	751	CATCGCAAGGTCTTCAAGCACTCGGATCAACGACGAAACATCCAGACCGCGCTCT	810
Db	320	GluGlyArgGlyGlyLeuIleArgArgValLeuArgArgProGlyArgProAlaPro	339
QY	811	CGCGTACCGCGAAACCGCTCGAACAGTCGCC-----	840
Db	340	GlyGlyProArgGlyHisArgGlnHisGlyProGlyIleArgHisLeuArgLeuLeu	359
QY	841	-----CCTCTACCAAAAAATACAGCGGGACGATCGCA	873
Db	360	ProGlyGlyArgAspHisProArgLeuProAlaProValArgThrSerGlyLysHisGly	379
QY	874	CTGGCA-----GAGCGTCCAACCCG	894
Db	380	GlnAlaGlyArgGlyLeuGlnGlnGlyThrGlyProValAlaArgGluGlyProArgAla	399
QY	895	CCTGATCAGCGACACCCCTGCGAAA-----AGGACTCGACCGCGCAGCGCGCAACC	945
Db	400	GlyValHisArgHisProAlaProGlyHisGlyArgSerArgGlyGlnProGlyArgThr	419

QY	946	GCCGATTCCGGGAGGCTGCAAGACGGCTCAAAACAGCCCGCAAAAAGCCTCTATCTGGT	1000
Db	420	Glu-----AlaP ^{ro} AlaGlyP ^{ro} AlaGluArgGlyP ^{ro} AlaGluArgGlnArgLeuGlnArg	437
QY	1006	TTCACCCTATTTCGTCCTCAAAATCCGGCAGACGCC-----ACTGCCAAACTGGT	1059
Db	438	ValP ^{ro} ArgP ^{ro} AlaTh ^{ro} AlaSerLeuGlnHisArgArgArgP ^{ro} ProAlaGlnArgGly	457
QY	1060	GCAGGACGGCATAGACGTTACCGTCCTGACCAACTCGCTACAGCGCACCGCTTGC	1119
Db	458	ArgArgArgHisArgGlyArgGlnArgGlu-----ArgArgAspArg	473
QY	1120	CGTCATTCCGGCTACGTCAAAATACCG-----AAAACCGCTGCTCAAGCCGGCAT	1170
Db	474	LeuP ^{ro} AlaArgArgP ^{ro} AspP ^{ro} ProGluGluArgGlyGlyAspArgA ^g His	493
QY	1171	CAAACTCTACGAGCTGCAACCCCAACCATGCGCTCCCGCCACAAAAGACAAAGCCCTGAC	1230
Db	494	HisLeuLeuHisGlnHisLeuGlnProGlnArgAspGlyArgArgProAlaGlyGlu	513
QY	1231	CGGCAGCTCCGTAACAGCTGTCATGCCAAACCTTCATTGTGGACGGCAAAACGCATCTT	1290
Db	514	GluGlyGlyGlyLysArgP ^{ro} AlaTh ^{ro} GlnAlaLeu-----GlyGluGlu-----	528
QY	1291	CATCGGCTCATCAACCTCGACCCCGCTCCGCACCGGCTCAATACCGAAATGGCGTCGT	1350
Db	529	-----PheAlaGlyAlaGlyLeuGlnGlyGly-----	537
QY	1351	CATCGAAAGCCCAAAATCCGACACAGATGGAGCGCACCTCGCGATACCCACACC---	1407
Db	538	HisArgLeuLeuGlnGlyArg ^g P ^{ro} AspP ^{ro} Leu---P ^{ro} ArgArgThrArgLeuArg	556
QY	1408	-----CGAATACCGCTACCGCGTTACCTCTCGACAAACACAAACCGCTCGCAATGGCA---	1458
Db	557	ProGlyArgLeuArgLeuHisLeuHisArgGlnLeuArgProAlaAlaGlyAlaAsp	576
QY	1459	CGATCCCGCCACCGCAAAAACCTACCGAACCGCAAGCCAAACTTTGGAAACGCAT	1518
Db	577	ArgGluSerHisProAlaGlyArg ^g P ^{ro} AspArg---ArgLeuGlyThrLeuArgGlnPro	595
QY	1519	CGCCGCAAAATCCT	1533
Db	596	GlnLeuArgArgPro	600

RESULT 14

US-09-134-001C-4273

; Sequence 4273, Application US/09134001C

; Patent No. 6380370

; GENERAL INFORMATION:

; APPLICANT: Lynn Doucette-Stamm et al

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS

; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: GTC-007

; CURRENT APPLICATION NUMBER: US/09/134,001C

; CURRENT FILING DATE: 1998-08-13

; PRIOR APPLICATION NUMBER: US 60/064,964

; PRIOR FILING DATE: 1997-11-08

; PRIOR APPLICATION NUMBER: US 60/055,779

; PRIOR FILING DATE: 1997-08-14

; NUMBER OF SEQ ID NOS: 5674

; SEQ ID NO 4273

; LENGTH: 495

; TYPE: PRT

; ORGANISM: Staphylococcus epidermidis

US-09-134-001C-4273

Alignment Scores:

Pred. No.:	8,84e-15	Length:	495
Score:	289.50	Matches:	105
Percent Similarity:	38.2%	Conservative:	68
Best Local Similarity:	23.2%	Mismatches:	151
Query Match:	10.2%	Indels:	129
DB:	2	Gaps:	15

US-10-665-990A-13 (1-1561) x US-09-134-001C-4273 (1-495)

```
QY 191 CATAACACGGGCTA-----TCCGACATCTACCTGCTCGACACGCCAC 235
   ::::|
Db 122 TyrAsnAsnAlaAlaPheLeuThrThrAspAsnLeuThrIleTyrThrAspGlyHis 141
   ::::|
QY 236 GAAGCCCTTCCGCGCGCCCTTATCGAATTCGCCGAACACAGCCTCGATTGCAA 295
   ::
Db 142 GlnLysPheAspAspLeuIleAsnAspIleArgHisAlaGlnSerTyrIleHisIleGln 161
   ::
QY 296 TACTACATTTGGCGGACACATTTCCGCGAGCTGCTTCAACCTCATGTACCTTGC 355
   ::
Db 162 TyrTyrIleIleHisSerAspAsnLeuGlyLysGlnLeuLeuIleGlnLeuLysLys 181
   ::
QY 356 GCAGAACCGCGGTGCGGTACGCTGCTGTTGGACGACAAACACACGCGCGTTGGAC 415
   ::
Db 182 AlaGluGluGlyIleGluValLysMetLeuTyrAspAspMetGlySerArgAspLeu--- 200
   ::
QY 416 GATCTCTCTCGCCCTCGACAGCCATCCCAATATCGAAGTGGCTGTTCAACCCCTTC 475
   ::
Db 201 -----ArgLysLys 203
   ::
QY 476 GTCCTACGCAATGGCGGCACTCGGCTACCTGACCGAC-----TTCCCC----- 520
   ::
Db 204 AspLeuLysLysPheArgGlnLysGlyGlyHisAlaGluSerPhePheProSerLysLeu 223
   ::
QY 521 -----CGCCTCAACCGCGCATGCAACAACAAATCTTTACCCCGCACAC 565
   ::
Db 224 ProLeuIleAsnLeuArgMetAsnAsnArgAsnHisArgLysIleValIleAspGly 243
   ::
QY 566 CGGCGCACCATCTCGGCGGACGCAATATCGGCGGCAATCTTCAAGTCGGTAGGAC 625
   ::
Db 244 ThrIleGlyTyrValGlyGlyPheAsnValGlyAspGluTyr-----IleGlyLysSer 261
   ::
QY 626 ACCTGTTTCCGC-----GACCTGGACATCTCGCCACCGGACGCTGCTCGCGCAA 676
   ::
Db 262 LysLysPheGlyTyrTrpArgAspThrHisLeuArgIleLysGlyAspAlaValAsnAla 281
   ::
QY 677 GTATCGCACACTTGCACCGCTACTGGGCAAGCCATTCGCCCCAC---AACGCCACGGC 733
   ::
Db 282 LeuGlnLeuArgPheIleLeuAspTrpAsnSerGlnSerThrArgAspAsnLeuThrTyr 301
   ::
QY 734 ATCATCCGC-----ACGGCAACATCGGCNAGGCTCTTCAGCA 772
   ::
Db 302 GluSerArgTyrPheProAspValAspSerGlyThrIleGlyIleGlnIleAlaSer 321
   ::
QY 773 CTCGGATACACGACGAAACATCCAGACACGCGCTCTCGCTACCGGAAACCGTCGAA 832
   ::
Db 322 SerGlyProAspGlu----- 326
   ::
QY 833 CAGTCGCCCTCTACCAAAAAATACAGACGGGACGATCGACTGGCAGAGCGCTCAA--- 889
   ::
Db 327 -----AspTrpGluGlnIleLysTyr 333
   ::
QY 890 -----ACCGGCTGATCAGCAGACCCCTGCAAAAGGACTCGACCGCCACGCCCAA 943
   ::
Db 334 GlyTyrLeuLysMetIleSerAla----- 342
   ::
QY 944 CGCGCGATTCCGGGAGGCTGCAAGACGCGCTCAAAACAGCCGCTCAAAACAGCCGCTATCT 1003
   ::
Db 343 -----LysGluSerIleTyrIle 348
   ::
QY 1004 GTTTCAACCTATTTCTCTCTCAAAATCCGGCAGACAGCGACTGCGCAAAACCTGTGTGAC 1063
   ::
Db 349 GlnSerProTyrPheIleProAspGlnAlaPheLeuAspSerIleLysIleAlaAlaLeu 368
   ::
QY 1064 GACGCATAGAGTTTACCGTCTGACCAACTCGCTACAGGCGACGACGCTGCGCGCCGTC 1123
   ::
Db 369 GlyGlyValAspValAsnIleMetValProAsnLysArg-----AspHisProPheVal 386
   ::
QY 1124 CATTCGGGTACGTCAAATACCGAAAAACCGCTGCTCAAAAGCGGCACTCAAACTCTACGAG 1183
   ::
```

```
Db 387 TyrTrpAlaThrLeuLysAsnValAlaSerLeuLeuGluAlaGlyValAsnValTyrHis 406
QY 1184 CTGCAACCAACCATGCGCTCCCGCCACAAAAGAGCAAAAGCGCTGACGGAGCTCCGTA 1243
   ::
Db 407 Tyr-----AspAsnGlyPhe----- 411
   ::
QY 1244 ACCAGCCTGCGATCCCAAAACCTTCAATCTGTGAGCGCAACGCGATCTTCAATCGGCTCATTC 1303
   ::
Db 412 -----LeuHisSerLysThrLeuValIleAspAspGluValAlaSerValGlyThrAla 429
   ::
QY 1304 AACCTCGACCCCGCTTCGCGACGCTCAATACCGAAATCGGCGTCTGTCATCGAAAGCCCC 1363
   ::
Db 430 AsnMetAspAsnArgSerPheThrLeuAsnPheGluValAsnAlaPheIleTyrAspGlu 449
   ::
QY 1364 AAATCGAGAACAGATGAGCGCACCTCGCCGAT----- 1399
   ::
Db 450 GlyValAlaArgSerLeuLysGlnAlaPheIleAsnAspMetLysLeuSerAsnLysLeu 469
   ::
QY 1400 ACCACACCGGATACGCTACCGGTTACCTCGACAAA 1438
   ::
Db 470 ThrSerGluGluTyrAlaLysArgAsnLeuValLys 482
   ::
```

RESULT 15

```
US-09-252-991A-21881
; Sequence 21881, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 21881
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-21881
```

Alignment Scores:

Pred. No.:	1,028-14	Length:	720
Score:	289.50	Matches:	160
Percent Similarity:	34.1%	Conservative:	41
Best Local Similarity:	27.2%	Mismatches:	209
Query Match:	10.2%	Indels:	179
DB:	2	Gaps:	30

US-10-665-990A-13 (1-1561) x US-09-252-991A-21881 (1-720)

```
QY 112 AGAAGCGGCGGAAAGCGCTCAATTTCAATACTTCCAAACCTGCTCTCGCAACATCCT 171
   ::
Db 179 ArgAlaThrGlyGlnProArgArgGlnGlyHisArgArgGlnArgProAlaProHisArg 198
   ::
QY 172 GCAAAATCCGGCACACCCCTCTCAATAACACGGGGTATCCGACATCTACCTGCTCGACGCC 231
   ::
Db 199 AlaAlaProAlaGluProGlyGlnGlyGlyAlaGlyArg-----AlaArgArgPro 215
   ::
QY 232 CCACGAAGCCCTTGGCGCGCGCGCTTATCGAATCTGCGAACACAGCCTCGATT 291
   ::
Db 216 -----GluProGlyArgArgArg----- 221
   ::
QY 292 GCAATACATATTGGCGCAACGACATTTCCGCGAGGCTGCTGTTCAACCTCATGTACCT 351
   ::
Db 222 -----ThrValGlnArgArg---ArgArgSerAlaArgGlnProGluGlnPro 236
   ::
QY 352 TCCCGCAGAACCGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGT 411
   ::
Db 237 GluArgGluProArgArg-----ArgAlaGlyGlnPro 247
   ::
```



```
Db      71 SerLeuPro-----:::|||||
QY      577 -----ACTCGCGGACGCAATATCGG---CGACGAATA-----606
Db      83 HisGlyGluProSerArgProAlaProAlaArgCysArgArgAlaAlaValAspProGly 102
QY      607 -----CTTCAAAAGTCGTGAGGACACCGTTTTCGC 636
Db      103 AspAspGlyGlyAsnGlyAspLeuAlaGlyIleGlnProArg-----ProPheArg 119
QY      637 CGACCTTGA-----CATCTCGCCACCGCGAGCGTCTCGG 672
Db      120 ArgHisGlyProHisProAlaLeuGlyGlnProAspProHisHisAlaAspArg-----137
QY      673 CGAAGTATCGCAGCATTCGACCGGTACTCGGCAAGCCATTCCGCCCAACGCCACGCGG 732
Db      138 -----LeuLeuLeuLeuArgProValLeu-----ProTrpArgProAlaArgProAla 153
QY      733 CATCATCCGACGGCAACATCGGCAAGGG-----TCTTCAAGCACTCGGTATACAAGA 786
Db      154 HisProProAspHisGlyArgLeuGlyValAlaGlyAspArgArgLeuArg 173
QY      787 CGAAAC---ATCCAGACCGCTCT---GGCTACCGGAAACCGTCAACAGTC 837
Db      174 ArgHisLeuValAspHisArgProGlyArgAlaValPheArgGlyArgHisVal 193
QY      838 GCCCTCTACCAAAAATACAGAC-----GGGACGCATCGACTG 876
Db      194 ArgProValProArgArgProLeuProArgThrProArgAlaGlyAlaHisArg 213
QY      877 GCAGAGCGT-----CCAAACCCG-----894
Db      214 LeuHisArgAlaThrGlyGlnProAlaAlaGlyValLeuProAlaProGlyArgGln 233
QY      895 CTGTATCAGCGACCCCTGCAAAAGGACTCCAGCGCGCCGCGCAACCCCGATTGC 954
Db      234 ProGlyArgAlaHisProAlaGlnArgThrAlaArgArgProArgThrGlyThrAla 253
QY      955 CGGGAG-----GCTGCAAGCGCGCTCAACAGGCCGA-----AAAAAGGCTCTA 999
Db      254 ArgArgAlaAlaGlyArgArgSerAspProArgArgProValGlnArgArgVal 273
QY      1000 TCTGTTTCCCTATTTCGTCCCTACAAATCCGGCAGACGCACT---GGCAAAATC 1056
Db      274 AlaAla-----HisArgValProAlaProSer 283
QY      1057 GGTGCGAGGACGATAGAGTTTACCGT-----CCTGACCAACTCGCTACGCGACCGA 1110
Db      284 ThrProGlyArgArgArgHisArgArgHisProGluArgArgGlyAlaAlaAspArg 303
QY      1111 CGTTGCGCGCTCCATTTCGGGTACTGTAATACCGAAACCGCTGCTCAAAAGCGCGCAT 1170
Db      304 ArgGlyArgArgProTrpArg---ArgGlnProPro-----ValGlyHis 317
QY      1171 CAAACTCTACGAGCTGCACCCCAACATCGCTGCCGTCGCCGCCACAAAGACAAAGCGCTGAC 1230
Db      318 ArgPro---ProAlaGlyThrArgPro-----GlyGlyGlnAlaAlaProGly 332
QY      1231 CGGAGCTCCGTAACAGCTGCATGCTCAACAAACCTTCA-----TGTGAGCGGAAACG 1284
Db      333 GlyThrGlyArgGlnGlyGlyAlaValValProAlaHisArgProAlaGlyArgArgSer 352
QY      1285 CATCTTTCATCGGCTCATTCACCTCGACCCCGGTTCCGACCGCT-----1329
Db      353 GlyArgAlaValValAlaGlyArgProProAlaArgLeuLeuGlyGlyAlaValAla 372
QY      1330 -----CAATACCGAAATGGCGTCT 1350
Db      373 AlaGlyArgHisLeuProLeuArgProValAlaGlyHisAlaHisArgAlaHisArgGly 392
QY      1351 CATCGAAAGCCC---CAAAATCGCAGAACGATGGAGCGCACCTCGC-----1395
|||||:::|||||
```

```
Db      393 HisArgGlnProAlaGlnAlaArgProAlaAspHisAlaArgAlaArgAlaGlyArgAla 412
QY      1396 -----CGATACCACACCGAATACGCTACCGCTTACCCTCGACAAACACACCGCCT 1449
Db      413 GluProHisArgHisGlyAsnLeuArgGlnAspArgHisProHisArgArgProAsp 432
QY      1450 GCAATGGCAGCATCCGCCACCGCAAAACCTACCCGAACGAAACCGAAGCCAACTTTG 1509
Db      433 AlaGlnAlaGlyAla-----ThrProAlaArgSerArgGlySerLeu 447
QY      1510 GAAACGCATCGCGC-----AAAAATCTATCCCTGCTGCCCATCGA 1551
Db      448 ProGlyProArgArgArgProArgGluProPheArgThrProHisArg 463

RESULT 17
US-09-252-991A-31760
; Sequence 31760, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 31760
; LENGTH: 1706
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-31760

Alignment Scores:
Pred. No.:      2,26e-14      Length:      1706
Score:          287.00      Matches:     163
Percent Similarity: 32.1%      Conservative: 52
Best Local Similarity: 24.3%      Mismatches: 207
Query Match:     10.1%      Indels:      248
DB:              2          Gaps:         31

US-10-665-990A-13 (1-1561) x US-09-252-991A-31760 (1-1706)
QY      199 CGGGCTATCCGACATCTACTCTCGACGACCCCAAGCCCTTGGCGCGCGCGCGC 258
Db      998 ArgLeuLeuLeuHisLeuArgProGlnArgProGlyGlnLeuProAlaGlnLeuArg 1017
QY      259 CCTTATCGAATCTGCGGAACACAGCTCGATTGCAATATCTACATTTGGCGCAACGACAT 318
Db      1018 AlaArgAlaAspArgArgAlaPro-----ProGlnArgGly 1030
QY      319 TTCGCGCAGGCTGCTTCAA-----CCTCATGTACCTTGC-----CGCAGAAGC 363
Db      1031 GlyAspProArgAlaAlaGlnGlyGlyAspHisLeuAspAlaAspArgHisLeu 1050
QY      364 CGCGCTGCGCGTACCTGCTGTTGGACGACAAACACACGCGCGGTTGGACGATCTCT 423
Db      1051 ArgHisAlaArgHisProAlaLeuAlaGluGlnProAlaArgLeuProAlaArgLeuPro 1070
QY      424 GCT---CGCCCTCGACAGCATCTCCCAATATCGAAGTGGCGCTTCAACCCCTTCGTCCT 480
Db      1071 GlyIleArgProArgThrValArg-----ArgValArgArgTyrArgPro 1086
QY      481 ACGCAATGGCGCGCACTCGGCTACTGACCGACTTCCCGCCCTCAACCGCGCATGA 540
Db      1087 ThrAlaArgAlaProAlaHisArgProGlnArgProAlaProProAlaLeuPro---Ala 1105
QY      541 CAACAAATCTTTTACCGCGCAACACCGGCCACCAT-----ACTCGCGCG 585
Db      1106 ThrArgArgLeuGlnArgAspArgArgValHisProLeuAlaGlyLeuValArgArg 1125
|||||:::|||||
```



```
Db 134 SerThrAlaThrAla-----ThrArgThrTrpProThrAlaSerArg 147
QY 522 GCTTCAACC-----GCGCATGCAACAACAAATCTTTACCGCGCACAAACCGCGCCACCA 575
Db 148 AlaSerThrProGlyAlaCysSerThrGlySerProProSerAsnCysAlaArgPro 167
QY 576 TACTCGGCGGAGCGRATA---TCGGGACGAAATACTTCAAGTCGGTGAGGACACCGTTT 632
Db 168 SerAlaArgProSerThrSerSerThrLysArgSerAlaSerAlaArg-HisArg-- 186
QY 633 TCGCCGACCTGACATCTTCGCGACCGCGAGGTGTCGGCGAAGTATCGCACGACTTCG 692
Db 187 ArgaspProGlyArgProAlaHisArgArg----- 197
QY 693 ACCGCTACTGGGCAAGCCATTCCGCCCAACACGCCGCGCATCATTCGCGACGGCAACA 752
Db 198 ----- 792
QY 753 TCGGCAAGGTCTTCAAGACTCGGATACACAGCAAAAC----- 792
Db 203 oArgArgGlyAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 223
QY 793 -ATCCAGACACGC-----GCTCTGCGCTACCGCGAAA 824
Db 223 uGlnArgArgProGlyGlnHisGlyProProProLeuAlaProGlyAlaProGlyVa 243
QY 825 CGCTCGAACAGTCGCCCCCTCTACCAAAA-----AACCGCGCTGATCAGGACA 908
Db 243 lLeuLeuGluAspAlaProLeuAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 283
QY 909 CCCTGCAAAAGACTCGACCGCGA-----CGCGCAAAACCGCGGATTG 953
Db 283 pProGlyArgArgProArgGlyArgProAlaArgLeuGlyProProCysLeuAlaAspPr 303
QY 954 CGGGAG-----GCTGCAAGACGCGTCAACAGCCCGCGGAAAAAAGCGTCT 998
Db 303 oArgArgArgLeuGluArgAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 323
QY 999 ATCTGGTTTC-----ACCTATTTCGCTCCTACAAAATCCGGCACAGAG 1043
Db 323 oGlyGlyHisProArgProAlaGlyAlaLeuPheArgArgArgArgAlaAlaAlaAla 343
QY 1044 CACTGCAAAACTGGTGACGAGCGCATAGACTTACCGTCTGACCAACTCGCTACAGG 1103
Db 343 gProGly-----GlyArg-----Gly 348
QY 1104 CGACCGAAGTTCGCGCGTCCATTCCGGCTACGTCAATACCGAATAACCGCTGCTCAAG 1163
Db 348 yGlnAlaAlaLeuArgLeuProLeuGlnArg-----LeuProArgThr-----Ar 364
QY 1164 CGGCATCAAACTACGAGCTGCAACCAACATGCCCTCCCGCCCAAAAAGACAAG 1223
Db 364 gArgValProProAlaProGlyArgArgGlnProTrpHisProAlaGlyAlaGlyGlyPr 384
QY 1224 GCTGAC-----CGGCAGCTCGTAACCGAGCTGCATGC-----CAAAACCTTCA 1268
Db 384 oAlaAspLeuArgArgGlyGlyArgSerArgProAlaAlaAlaArgGlyLeuGlnArgLeuHi 404
QY 1269 TTGTGA-----CGCAAAACGCATCTTCAT-----CGGCT 1298
Db 404 sProGlyThrLeuArgArgTrpProArgGlnArgLeuHisHisProArgArgGlyArgGl 424
QY 1299 CATTCAACCTCCAGCCCCCTTC---CGCAGGCTCAATACCGAATGGCGCTCGTCATCG 1355
Db 424 yAlaAlaProArgProGlyValProArgThrAlaAlaAlaAlaAlaAlaAlaAlaAla 444
QY 1356 AAAGCCCCAAAATCGCAGAACAA-----GATGAGGCGCACCTTCGCGCATACACAC 1406
Db 444 oLeuProProValGlyArgThrAlaAlaGlyArgSerAlaGlnProAlaAlaGlyArgTh 464
```

```
QY 1407 CCGAATACGCTACCGCTTACCTCGA-----CAAACACAACGCGCTGCAATCGCACG 1460
Db 464 rGlyAlaArgSerProGlyArgProArgGlyLeuAlaGlySerAlaProAlaMetSerAr 484
QY 1461 A-----TC 1463
Db 484 gArgGlnThrCysSerLeuLeuLeuAlaPheGlyLeuPheTyrLeuValProLeuSe 504
QY 1464 CCGCCACCCGAAAAAACCTACCGAACGACGAAACCGCA 1497
Db 504 rAshHisGlyLeuTrpIleProAspGluThrArg 515

RESULT 21
US-09-252-991A-20186
; Sequence 20186, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252.991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 20186
; LENGTH: 1098
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-20186

Alignment Scores:
Pred. No.: 6,898-14 Length: 1098
Score: 280.00 Matches: 171
Percent Similarity: 35.7% Conservative: 43
Best Local Similarity: 28.5% Mismatches: 195
Query Match: 9.8% Indels: 190
DB: 2 Gaps: 39

US-10-665-990A-13 (1-1561) x US-09-252-991A-20186 (1-1098)
QY 136 CAATACTTCCAAACCTGCTCTCTGGACAAACATCTGCAAAATCCGGCACACCCCTCATAA 195
Db 3 ArgTyrPheArgAlaProGluProPheSerProArgValThrGluGluSerGln 22
QY 196 CACCGGTATCCGA-----CATCTACTGCTCGACGACCC----- 231
Db 23 ProArgLysIleHisTrpSerValProHisValProValThrArgProValAlaAlaGln 42
QY 232 -----CCACGAAGCCTTGGCGCCCGCGC-----CGCCTTATCGAATC 270
Db 43 ProArgProAlaProArgLeuProArgSerArgProGlyPheArgArgGlnArgLeu 62
QY 271 TGCCGAACACAGCCT---CGATT---GCAATPACTATTTTGGCGAACGACATTTCCCG 324
Db 63 ArgArgArgAspProHisArgTyrProProValLeuProAlaAlaArgAlaGly 82
QY 325 CAGGCTGCTGTTCACCTCATGTACTTCCCGCAGAACG----- 363
Db 83 AlaGlyAlaGluProProArgProProGlyArgArgAlaAspArgLeuGlnProGlyThr 102
QY 364 ---CGCGCTGCGCGTACGCTCTGTGGACCAACACACGCGCGGTGGAGCATCT 420
Db 103 AspArgArgAlaAlaGlnProGlyAlaGlyArgArgValHisPro-----GlyGlySer 120
QY 421 CTGCTCGCCTCGACGCATCCCAATATCCAAATATCGCGCTGTTCACCCCTTCGT--- 477
Db 121 ProGlyArgProAlaArg-----ArgLeuArgProGlyGlyAlaAlaArgProArgArg 138
```

```
QY 478 -----CCTACGAAATGGCGCGCCTCGCTACCTGGCTACCGCGACTTCCCGCGCTCAA 528
Db 139 ArgLeuTyrProGlyGlyAlaGlyArg-----ProGlyGln 152
QY 529 CGCGCGCATGCAACAATCTTTTACCGCGACAAACCGCGCCACCATCTCGCGGACG 588
Db 153 ProAlaGlyAla-----HisArgGlyArgLeuAlaGlyArg 166
QY 589 CAATATCGCGACGAATCTTCAAGTCGGTCCAGGACACCGCTTTT----- 633
Db 167 GlnArgArgArgLeuArgAlaSerArgCysAlaArgArgGlyProProArgAlaVal 186
QY 634 -----CGCGCACTCGGCATCCT----- 651
Db 187 ProGlyProArgArgGlyGlnHisProArgSerAlaGlyAlaHisSerArgGlyGlnArg 206
QY 652 -----CGCCACCGGAGCGTCT-----CGCGAAGTATCGCA 684
Db 207 AlaAlaGlnArgHisArgGlnProArgProGlyIleGluLeuArgHisSer----- 224
QY 685 CGACTTCGACCGCTACTCGGC----- 705
Db 225 ArgProGlnProAlaProGlyValAlaLeuAspGlyAlaAspGlyArgTyrProGlyAla 244
QY 706 ---AAGCCATTCCGCCCAACAGC-----CACGCGCATCAT--- 738
Db 245 LeuArgProLeuArgProAlaThrValValAlaGlyAlaGlyValHisArgGlnHisGly 264
QY 739 -----CCGACGGGCAACATCGCAAGGTTCTCAAGCACTCGGATACAA 783
Db 265 ArgGlyGlyArgGlyProArg-----ArgArgGlyAlaLeuArgAlaAlaGluArg 282
QY 784 CGAGCAACATCCGACACGCGCTCTCGCTACCGCAACCGTCCGACAGTCCGCCCT 843
Db 283 ArgArgHisArgGlnLeuArgAspProGlyAspProArgGlyLeuArgHisGlnAla--- 301
QY 844 CTACCAAAAAATACAGCGGACGATCGACTGGCAGACGCTCCAAACCGCCTGATCAG 903
Db 302 -----ArgArgAlaGlnArgThrGlnPro-----Gln 310
QY 904 CGACACCCCTCAAAGACTCGACCGGACCG---CGCAACCGCGGATTCGCGGAG 960
Db 311 LeuGlnPro-----GlyArgProGluAspHisProGlnArgAlaAsp----- 324
QY 961 GCTCAAGACGCGCTCAAACAGCCGAAAAAGCGTCTATCTGTTTCAACCTATTTCGT 1020
Db 325 -----ArgArgHisArgArgGlnArgProArg-----Arg 335
QY 1021 CCTACAAAATCCGCGCACAGCACTGGCAAAACTGGTCGACGCGCATAGACGTTAC 1080
Db 336 ProAlaLeuLeuArgHisProArgArgArgLeuAlaArgAlaGlnArgTyrAlaAspArg 355
QY 1081 CGTCTGTAC-----CAACTCGCTACA-----GGCGAC--- 1107
Db 356 ArgProAspProGlnGlyProLeuProAlaGlnArgArgThrHisValPheGlyAspAsp 375
QY 1108 -----CGACGTTGCGCGCTCCATTTCGCGCTACGTCATCAATACCG 1146
Db 376 ProValLeuArgArgArgGlyArgHisAlaArgArgPro---ArgHisArgGlyLeuPro 394
QY 1147 AAAACCGCTGCTAAAGCGCGCATCAAACTCTACGAGCTGCAACCCCAACCATGCGTCCC 1206
Db 395 ArgArgProValProVal---AspProSerLeuArgGlnValLeuGlyProProTyrPro 413
QY 1207 -----CGCCNAAAGAACAGCCTGACCGGAG----- 1236
Db 414 GlyGlnArgGlnLeuArgGileHisProGlnArgGlnProGluAlaGlnArgHisArgLeu 433
QY 1237 CTCGCTAACGCTGATGATGCCAAACCTTCATTGTGACGGCAACCGCAT-----CTT 1290
Db 434 LeuHisGlnAspProAlaGlnArgLeuProArgProGlyProGlnProHisProValAla 453
QY 1291 CATCGGCTCATTAACCTCGACCC-----CCGTTCCGCGCGGCTCAATACCGAAATGGG 1344
```

```
Db 454 AlaArgValLeuGlyThrArgProGlyAsnProPheGlnProGly----LeuArgAlaGly 472
QY 1345 CGTCGT-----CATCGAAAGCCCAAAATCGCAACACAGATGGA----- 1383
Db 473 ArgLysSerProArgSerGlyHisArgProProLeuArgGlnArgSerGlnProArgAla 492
QY 1384 -----CGGCACCTCCCGCA-----TACCACACCCGNAATACGCTACCG 1422
Db 493 AlaLeuLeuAspProArgArgGlnArgProAlaThrGlnHisArgGlnProGlnArgPro 512
QY 1423 CGTTTACCTTCGACAAACACCAACCGCTGCAATGCGACGATCCCGC----- 1467
Db 513 ArgHisProArgGlnHisArgSerArgValLeuHisArgArgSerHisArgHisArg 532
QY 1468 -----CACCCGAAACCTACCCGAAACCCGAAACCCGAAACCTTTGGAA 1512
Db 533 GlnLeuAspHisHisProArgHisProLeuArgGluAspArgPheArgThrGluGlu 551

RESULT 22
US-09-902-540-16463
; Sequence 16463, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; PRIORITY FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIORITY FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 16463
; LENGTH: 420
; TYPE: PRT
; ORGANISM: Myxococcus xanthus
; US-09-902-540-16463

Alignment Scores:
Pred. No.: 1,19e-13 Length: 420
Score: 275.00 Matches: 114
Percent Similarity: 36.8% Conservative: 65
Best Local Similarity: 23.5% Mismatches: 179
Query Match: 9.6% Indels: 128
DB: 2 Gaps: 15

US-10-665-990A-13 (1-1561) x US-09-902-540-16463 (1-420)
QY 149 CTTGTCCTCTCGCAACATCTCTGCAAAATCGGCGACACCCCTCATACACGCGGCTA--- 205
Db 38 ProGlyValSerGlyGlyLeuLeuThrArgTyrTyrLeuProArgArgHisGlyValVal 57
QY 206 -----TCCGACATCTACCTGCTCGACGACCCACGAGCCCTTTCGCGCGCGCGCC 259
Db 58 GlnGlyAsnAlaCysGlnLeuLeuArgAspGlyValGluAlaTyrProAlaMetLeuGlu 77
QY 260 CTTTACGAATCTCCGAACACAGCCTCGATTTCGAATCTACATTTTGGCGCAACGACATT 319
Db 78 AlaileArgGlyAlaArgArgTyrValArgMetGluThrTyrMetPheValSerAspAla 97
QY 320 TCCGCGAGGCTGCTGTTCAACCTCTGCTGCGCAGAACGCGCGCTGCGGTACCG 379
Db 98 ValGlyLeuLeuPheGlyGlnAlaLeuAlaGluAlaGluArgGlyValHisValLys 117
QY 380 CTCGCTGTTCGAC-----GACAACAAC 400
Db 118 ValLeuTyrAspAlaValGlySerTrpThrSerArgArgSerPhePheAlaGlyLeuArg 137
QY 401 ACGCGCGGGTTGGACGATCTCTCTGCTCGCCTCGACAGCCATCCCAATATCGAAGTCGCG 460
Db 401 ACGCGCGGGTTGGACGATCTCTCTGCTCGCCTCGACAGCCATCCCAATATCGAAGTCGCG
```

```
Db      138 AlaArgGlyValasp-----lIeArg 144
Qy      461 CTGTTCAACCCCTTGGTCTTACGCAAAANGCGCGACTCGGTACTGTACGTACCGACTTCCCC 520
Db      145 AlaPheLysProPheSerLeuSer-----ArgGlyLeuArgHisLeuLeu----- 159
Qy      521 CGCTCAACCGCGCGATGCACAAACAAATCTTTACCGCGCGACAAACCGCGCCACCATCTC 580
Db      160 -----ArgArgAspHisArgLysIleLeuValValAspGlyGlyValAlaPheThr 176
Qy      581 GCGGACGCAATATATGCGGACGAAATACTTTCAAAGTCGGTGAGGACACCGTTTTTCGCCGAC 640
Db      177 GlyGlyValAsnIleSerAlaHisIleProAlaProAlaGluMetGlyAlaIleIlePheArgasp 196
Qy      641 CTGGACATCTCTCCGACCGCGAGCGTGTGGCGAAGTATCGACGACTTCGACCGCTAC 700
Db      197 AspValLeuArgIleGluGlyProAlaValHisGluLeuGluArgCysPheSerAlaThr 216
Qy      701 TGGCGAAGCCATTCGCCCCACAAACGCGCAGCATATCCGACGGCGGCAACATCGGCAAG 760
Db      217 Trp-----ArgMetMetPheGlnGlyArgPhe----- 225
Qy      761 GGTCTTCAAGCACTCGGATACAAACGACGAAACATCCAGACACGCGTCTCGGTACCGC 820
Db      226 -----HisArgLeuThrArgArgLeu 232
Qy      821 GAAACCGTGCAGACGTCGCCCTCTTACCAAAAATAACAGCGGACGCGTACGTGGCAG 880
Db      233 GluArgLeuArgAsnProPro----- 239
Qy      881 AGGCTCAAAACCGCGCTGATACGACACCCCTGCNAAGGA-----CTC 925
Db      240 -----ProArgArgGlyAlaValGlyLeuValVal 249
Qy      926 GACCGGACCGCGCAAAACCGCGGATTGCGGGAGGCTGCAAGCGCGTCAAAACAGCC 985
Db      250 LeuSerSerArgArgSer-----IleHisArgAlaTyrLeuHisAlaIleIleArgAla 267
Qy      986 GAAAAAGCGTATCTGTGTTTACCCCTATTTCCTTACAAATTCGGGACAGACGCA 1045
Db      268 ArgArgSerValLeuValAlaAlaIlePheIleProAspArgMetValMetAla 287
Qy      1046 CTGGCAAAACTGGTCAGACGCGCATAGAGTTACCGTCTGACCAACTCGTACAGGG 1105
Db      288 LeuArgGluAlaAlaArgArgGlyValGluValHisLeuLeuLeuAsnAla-----Arg 305
Qy      1106 ACCGAGCTTTCGCGCGCTCCATTCCGGCTACGTCAAATACCGAAACCGCTGCTCAAAGCC 1165
Db      306 SerAspHisProIleLeuGluPheMetAlaArgAlaPheTyrGluArgLeuLeuGlyAla 325
Qy      1166 GGCATCAAACTTACGAGTGCAAACCCAAACCATGCGCGTCCCGCGCCAAAGACAAAGCC 1225
Db      326 GlyValArgIlePheGluTrpGln-----ArgGly 335
Qy      1226 CTGACCGGAGCTCCGTACACAGCTGCATGCGCAAAACCTTCATTGTTGACGCGCAACGC 1285
Db      336 Val-----LeuHisAlaLysThrAlaValValAspGlyValTrp 348
Qy      1286 ATCTTCATCGGCTCATTAACCTCGACCCCGCTTCGCGACGGCTCAATACCGAAATGGGC 1345
Db      349 GlyThrIleGlySerPheAsnLeuGluArgLeuSerLeuAlaPheAsnHisGluValAsn 368
Qy      1346 GTCGTCATCAAGAGCCCAAAATTCGCAGAACAGATGGAGCGGACCCCTCCCGCATACCACA 1405
Db      369 AlaValPheAlaAspProArgLeuGlyGlnGlnLeuGluAspSerPheArgGlyAspCys 388
Qy      1406 CCGGAATACGCTACCGCTTACCTTCGACAAACACACCGCTGCAATGGCACGATCCC 1465
Db      389 GlyAsp---CysArgGluValThrLeuAlaGluPheArgArg----- 401
Qy      1466 GCCACCCGAAAAACCTTACCCGACGAAACCGAAGCCAAACTTTTGGAAACGCATCGCGCA 1525
Db      402 -----ArgProLeuTrpGlnLysLeuLeuGlu 410
```

```
Qy      1526 AAAATCCTATCCTCGCTG 1543
Db      411 ArgAlaLeuSerLeuLeu 416

RESULT 23
US-09-976-594-615
; Sequence 615, Application US/09976594
; Patent No. 6673549
; GENERAL INFORMATION:
; APPLICANT: Furness, Michael
; APPLICANT: Buchbinder, Jenny
; TITLE OF INVENTION: GENES EXPRESSED IN C3A LIVER CELL CULTURES TREATED WITH STEROIDS
; FILE REFERENCE: PA-0041 US
; CURRENT APPLICATION NUMBER: US/09/976,594
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: 60/240,409
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 1143
; SOFTWARE: PERL Program
; SEQ ID NO 615
; LENGTH: 904
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. 6673549 2789525CD1
US-09-976-594-615

Alignment Scores:
Pred. No.: 1,76e-13 Length: 904
Score: 274.50 Matches: 133
Percent Similarity: 37.7% Conservative: 60
Best Local Similarity: 26.0% Mismatches: 204
Query Match: 9.6% Indels: 115
DB: 2 Gaps: 23

US-10-665-990A-13 (1-1561) x US-09-976-594-615 (1-904)
Qy      217 CTGTCTGACGACCCCGACGAGCCCTTTCGCGCGCGCGCCCTTATCGAATCTGCCGA 276
Db      172 ProArgArgArgLySerArgSerProArgArgArgSerProValArgArg 191
Qy      277 ACACAGCCTCGATTGTCATCTACTACTATTGGCGCAACGACATTTCCGCGAGCGTCTGTT 336
Db      192 GluArgLysArg-----SerHisSerArgSer----- 200
Qy      337 CAACCTCATGTACCTTTCGCGAGAACGCGG---CGTGGCGGTACGCTGCTGTGGACGA 393
Db      201 -----ProArgHisArgThrLysSerArgSerProSerProAlaProGluLys 216
Qy      394 CAACAACACGCGCGGTGGACGATCTCTGCTCGCCCTCGACAGCCATCCCAATATCGA 453
Db      217 LysGlu-----LysThrProGluLeuProGluProSerValLysValLys 231
Qy      454 AGTGGCCTGTTCAACCCCTTCGTCTCTAGC---CAAATGGCGCGCACTCGGCTACTGAC 510
Db      232 GluProSerValGlnGluAlaThrSerThrSerAspIleLeuLysValProLysProGlu 251
Qy      511 CGACTTCCCGCTCAACCGCGCATGCACAAACAAATCTTTACGCGCAACACCGCGC 570
Db      252 ProIleProGluProLysGluProSerProGluLysAsnSerLysLysGluLys 271
Qy      571 CACCATACTCGG---CGGACGCAATATCGCGCAAGAACTACTTCAAAGTCGGTGAGACAC 627
Db      272 GluLysThrArgProArgSerArgSerLys-----SerArgSerArgThr 288
Qy      628 CGTTTTCCCGACCTGGACATCTCGCCACCGGACGCT----- 666
Db      289 ArgSerArgSerProSerHisThrArgProArgArgHisArgSerArgSerArgSer 308
Qy      667 -----CGTGGCGGAAGTATCGACGACTTCGACCGCTACTGGGCAAGCCATTCGCC 717
```


Qy	1171	CAAACTCTACGAGCTGCAACC-----CAACATCGCGTCCCGCCAC	1212
		::: ::	
Db	591	ArgLeuProAlaValAlaGlyProGlyGluLeuHisLysProGlyArgLeuAlaAsp	610
		::: ::	
Qy	1213	AAAAGACAAAGGCGTGAC-----	1230
Db	611	HisArgGlnArgProAlaGlyGlnLeuAlaGlyGlyGlnAlaAlaArgArgArgArg	630
Qy	1231	-----CGGAGCTCCGTAAACACCGCTCATGCCAAAAACCTT-----	1266
		:: ::	
Db	631	ProGlyAlaGlyLysIleAspGlnLeuGlnProArgProGlyAlaArgSerLeuAlaGly	650
		:: ::	
Qy	1267	---CATGTGGACGGCAACGCATCTTCATCGCTCATTTCAACTCGACCCCG-----	1317
		::	
Db	651	ProGluProGlyHisArgArgLeuSerHisArg-----AspProArgProHisArgArg	668
		::	
Qy	1318	-----TTCCGACCGCTCAATACCGAAATGGGGTGGTGGTCAATGAAAAGCCCCAA	1365
Db	669	ArgCysHisLeuPheArgThrAlaGlyAspArgCysAlaAlaArgArgArgAspGlnHis	688
Qy	1366	AATCGCAGAACAGAT--GGAGCGCACCTTCGCGCATACCACCCGAATACGCTACCG	1422
Db	689	ProGlyArgThrAspProGlyGlnHisPro-----LeuHisAspGlnArgArgArgHis	706
Qy	1423	CGTTACCCTCGA-----CAAAACACAAACCGCTGCAATGGCAGCATCCCGCCACCCG	1473
Db	707	ProTyrProArgProArgProHisArgGlnLeuProAspProAlaArgArgValGlyPro	726
Qy	1474	AAAAACCTACCGAAGCAACCGAAGCCAACTTTGGAAACGCATCCGCGCAAAAATCCT	1533
		:: ::	
Db	727	AspArgLeuAlaProGlyArgGlnArgGlnSerHisGluAlaGlyGlnGluSerSerArg	746
		:: ::	
Qy	1534	ATC-----CCTGCTGCCATCGAAGG	1554
Db	747	AlaGlnArgProAlaAlaGlnArg	755

```

US-09-252-991A-26099
; Sequence 26099, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIORITY FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIORITY FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIORITY FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 26099
; LENGTH: 863
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-26099

```

Alignment Scores:					
Pred. No.:	3,95e-13	Length:	863		
Score:	270.00	Matches:	174		
Percent Similarity:	32.0%	Conservative:	36		
Best Local Similarity:	26.5%	Mismatches:	205		
Query Match:	9.5%	Indels:	242		
DB:	2	Gaps:	33		
US-10-665-990A-13 (1-1561) x US-09-252-991A-26099 (1-863)					
Qy	208	CGACATCTACTGTCGCAGCACCCCA	CGAACGCCCTTGCCGGCCGCCCTTATCGA	267	
Db	45	ArgArgProSerGlnArgArgGlyAspProArgLeuArgLeuProLeuArg		64	

QY	268	ATTCGCGAACACAGCCT	-----CGATTTCGAATACTACAT	-----	303
DB	65	GluArgArgIleAerProCysLeuArgSerArgAspCysPheHisAerProGlyCys	 ::: 		84
QY	304	-----	-----TTGGCGCAACGACATTTC	-----	327
DB	85	GlySerHisProAspGlyGlnGlnAlaProValGluAspArgHisAlaAerArgArg	::: ::: 		104
QY	328	GCTGCTGTTTCAACCTCATGTACCTTGCAGCAACGCGCGTGCCTACGCTCGCTGTT	 ::: 		387
DB	105	GlyAlaValHisProArgLeuAerArgArgThrGlyArg--ArgAspProAlaAla	 ::: 		123
QY	388	-----	-----GGACGACAA	-----	396
DB	124	ArgGlyAerThrHisAerLeuProThrAspGlnGlyGlnAerArgAerArgPro			143
QY	397	CAACACGCG	-----	-----	405
DB	144	ArgHisAlaProGlyGlyGlyThrArgProThrAlaGlyProValAlaHisAerThrLeu	::: 		163
QY	406	---CGGTTTGCAGCATCTCTGCTCGCCCTCGACAGCATCCCAATATCAAGTCGCCT	 ::: 		462
DB	164	GlySerAlaGlyAerLeuAerGlnAerArgThrAspProGlyGlyAerAlaAer	 ::: 		183
QY	463	GTT-----	-----CAACCCCTTCGCTCTACGCAAAATGGCGCG		495
DB	184	AlaProCysGlyAspProGlyValAerArgProProArgGlnProAlaLeuProArg	 ::: 		203
QY	496	ACTCGGCTACTGACGACATTCCTCCCGCTCAACCGCGCATGCACAAATC	 ::: 		549
DB	204	ThrArgLeuLeuGlyAlaAlaProProProGlyGlyHisAerGlySerSerLeuProGly	 ::: 		223
QY	550	---CTTTACCGCGACAAACCGCGCCACCATACTCGCGGACGCAATATCGCGACGAATA	 		606
DB	224	AlaAerCysArgProAlaSerAerHis		-----	232
QY	607	CTTCAAAAGTCGTGAGGACACCGTTTTTCGCCGACCTGGACATCTCGCCACCGCGACGT	 		666
DB	233	-----GlyAerGlyGlyGlyGluGlyAerGlyPheGlyGlnLeuAerGlyAerHis	 		250
QY	667	CGTCGGCGAAGTATCGACAGACTTCGACCGCTACTG-----GGCAGCCATTTCGCCCA	 		720
DB	251	ArgAerValSerSerAerAerGlyTrpProLeuLeuLeuProGlyAerGlyAerProPro	 		270
QY	721	-----	-----CAACGCCACGCGCATCATCCG	-----	747
DB	271	AlaGlyAerThrProGlyHisAerGlyThrGlyHisAerProGlyProGlyGlyLeuAlaAer	 		290
QY	748	CAACATCGGCAAGGCTTCAAGCACTCGGATACAA	 	-----	783
DB	291	ProAerArgAerGlyProSerAlaAlaAlaAlaGlyThrGlyGlyThrAlaAerAla	 		310
QY	784	-----CGAGAAACATCCAGACACGCGCTCTCGCTACCGCGA	 		822
DB	311	CysHisAerGlyProProLeuAerAerGlyAerAerSerArgLeuProAlaAlaAerAer	 		330
QY	823	AAC-----	-----CGTCAACAGTCGCGCCCTCTACCAAAA		852
DB	331	HisGlyAlaAlaLeuGlyProGlyThrAlaAlaAerArgThrHisAerProTrpProGly	 		350
QY	853	AATACAGACGGACGCATCGACTGGCAGAGCGTCCAAACCGCGCTGATCAGCAGCACCCC	 		912
DB	351	ArgGlyAlaGlyAerAlaLeuLeu--AerProAer--AlaGlyGlnAlaAerAer	 		368
QY	913	TGCAAAAGGACTCGACCGGACCGCGCAACCCCGGATTGCCGGGAGGTGCAAGACGC	 		972
DB	369	LeuAerGlyAerProAerAerSerProAlaGlnAlaAerProAerAerGlyAlaLeuAer	 		388
QY	973	GCTCAACACGCCGAAAAAG-----	-----CGTCTACTCTGGTTTC		1008
DB	389	AlaAlaGlyAerAerGlnProAlaValProAerGlnProProGlyAlaSerGlyLeu	 		408
QY	1009	ACCCTATTTCGTCCTCAAAATCCGCGACAGCGCACT-----	-----	-----	1047

RESULT 27

RESOL 2 /
US-09-252-991A-18427

: Sequence 18427, Application US/09252991A

Patent No. 6551795

: GENERAL INFORMATION:

APPLICANT: Marc J. Rubenfield et al.

APPLICANT: M&C S. ANGENEN CO. INC.
 TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS

FILE REFERENCE: 107196-136

FILE REFERENCE: IO/190.130
CURRENT APPLICATION NUMBER: IIS/09/252.991A

;; CURRENT APPLICATION NUMBER: US/00-02-18
: CURRENT FILING DATE: 1999-02-18

; CURRENT FILING DATE: 1999-02-18
 ; PRIOR APPLICATION NUMBER: US 60/074 788

;; PRIOR APPLICATION NUMBER: US 60
: PRIOR FILING DATE: 1998-03-18

; PRIOR FILING DATE: 1998-02-18
 : PRIOR APPLICATION NUMBER: US 60/094 190

;; PRIOR APPLICATION NUMBER: US
: PRIOR FILING DATE: 1998-07-27

;; PRIOR FILING DATE: 1998-07-13
; NUMBER OF SEQ ID NOS: 33142

; NUMBER OF SEQ
SEQ ID NO 18437

; SEQ ID NO 184

US-10-665-990A-13 (1-1561) x US-09-252-991A-18427 (1-748)

Qy	961	GCTGCAAGACGGCGTCAAAACAGCCGCAAAAAGCGTCTATCTGGTTTCCACCTATTTCGT	1021
Db	295	-----LeuAspArg	297
Qy	1021	CCCTACAAATCCGG-----CACAGCGC-----ACTGGCAAAACTGGT	1059
Db	298	ProTrpArgLeuAlaGlnProGlyHisGlnArgGlyAlaAlaLeuAlaGlyGlnArgArg	317
Qy	1060	GCAGGACGGCAT-----AGACGTTACCGTCTCTACCAACTCGCTACAGCGCGCCGA	1110
Db	318	SerGlyLeuHisProArgProArgLeuTrpArgProGly-----AlaAspGlu	333
Qy	1111	CGTTGGCCGCGTCCATTCGGCTACGTCAATACCGAAACCCGCTGCTCAAAACCGCGCAT	1170
Db	334	ArgLeuArgHisProAlaGlyArgGlnValProGlnAlaProAlaArgGluGluLeu	353
Qy	1171	CAAACTCTA-----CGAGCTGCAACCCAA-----	1194
Db	354	GlnArgAlaGlnArgArgAlaLeuGlnArgProValAspProProAspGlnArgLeu	373
Qy	1195	---CCATGCGTCCCGCCACAAAGACAA-----AGCCCT	1227
Db	374	GluProPheArgGlnArgHisArgGluGlnProValValLeuHisGlnGlyProProPro	393
Qy	1228	GACCGCGACCTCCGTAAAC-----CAGCCTGCATGCCAAAACCTTCATTGTGGA	1275
Db	394	LeuAlaGlnArgArgAsnLeuArgMetGlyArgProAlaArgAlaValAlaGlnGly	413
Qy	1276	CGGCAAAACGATCTTCATCGG-----CTCATTCAACCT	1308
Db	414	LeuProAlaHisGlnProArgAlaGlyAlaAspArgProProAspLeuArgLeuPro	433
Qy	1309	CGACCCCGTTCGCGACGGCTCAATACGAAATGGCGGTGCTCATCGAAAGCCCCAAAAT	1368
Db	434	AlaArgProValArgProArgGlnProHisProGlyArgArgValGlnProHisPro	453
Qy	1369	CGCAGAACAGATGGAGCGACCCCTCGCCGATCCACACCGCAATACGCTACCGCGT---	1425
Db	454	LeuProProAspGlnGlnHisIleAlaLeuTyArgCysArgArgArgLeuHisArgPro	473
Qy	1425	-----	1425
Db	474	LeuAlaSerArgAlaArgLeuLeuArgLysProLeuAlaLeuProAlaAlaLeuAlaGln	493
Qy	1426	-----TACCTCTGCACAAAACACACCCGCTTGCATGCGCATCC-----	1464
Db	494	ProAspProHisLeuArgProValArgArgGluProProAlaAlaGlnArgAlaProVal	513
Qy	1465	-----CGCCACCGCAAAAACCTACCCGCAACGAAACCGGACCGCAAACTTTGGAA	1512
Db	514	AlaGlyAspArgGlyAlaProArgProGluProTyArgProArgArg-----Pro	530
Qy	1513	ACGCATCGCGC---AAAAATGCTATCTCCTGCTGCCCATCGA	1551
Db	531	AlaHisArgHisProGlnArgProGlnProAlaGlyArgGln	544

US-09-252-991A-19467

```

US-09-252-991A-1946 /
; Sequence 19467, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142

```

```
; SEQ ID NO 19467
; LENGTH: 470
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-19467

Alignment Scores:
Pred. No.: 3,76e-13 Length: 470
Score: 269.00 Matches: 140
Percent Similarity: 33.6% Conservative: 33
Best Local Similarity: 27.2% Mismatches: 192
Query Match: 9.4% Indels: 150
DB: 2 Gaps: 25

US-10-665-990A-13 (1-1561) x US-09-252-991A-19467 (1-470)

QY 76 TCTCTGTTTCATGTTCTTCATGTTGCCCTCCCTGGAAGAACGGACGAAAGCGTCATTT 135
DB 15 SerLeuTyrGlyPheGlnSerIleAlaProGlyAlaLeu----- 28
QY 136 CAATACTTCAAACCTGCTCTCTGGACAAACATCTTGCAAATCCGGACACACCCCTCATAA 195
DB 29 -----SerArgAlaProGlyArgGlyAlaGlyArgGlyIleHis----- 41
QY 196 CAACGGGCTATCGACATCTACTGCTCGACGACCCCGACGAGCCCTTGGCCGCGCGC 255
DB 42 -----ArgArgAlaGlyArgGlyAspProGlyArgProAla 52
QY 256 CGCCCTTATCGAATCTCGGACACACAGCTCGATTGTCATCTACTATTTGGCGCAACGA 315
DB 53 AspProAlaArgLeuGlyCysArgProGlnPro-----AlaThrAspArg 66
QY 316 CATTTCCGGCAGCGTCTGTTCAACCTCATGTACCTTGGCCGAGAACGGG----- 366
DB 67 ProPheHisGlnLeu-----TyrArgArgSerArgGlyGlyLeu 79
QY 367 -----CGTGGGTAGCTGCTGTTGGAGCAGACACACCGCGGGTT 411
DB 80 ArgArgCysProGlyArgArgGlyValProGlyAlaAspArgLeuGlnArgArgGlyLeu 99
QY 412 GGACGATCTCTGCTGCGCCT-----CGACAGCATCCCATATATCGAAGTGGCGTGT 465
DB 100 GlnArgHisArgLeuArgProProGlyThrAspProGlyGlnLeuAlaSerGlyHis 119
QY 466 CAACCCCTTCTGCTACGCAATGGCG-----CGC 495
DB 120 ArgAlaAlaAlaProAlaGlyThrAlaThrGlyAlaLeuAlaAlaGlnArgAlaAspArg 139
QY 496 ACTCGGCTACCTGACGAGCTTCCCGCTCAACCGCCGATGCAACAATACTCTTTTAC 555
DB 140 ThrThrAspProAlaArgArgProGluAspArgPheGlyGlyGlyGlyLeu----- 157
QY 556 CGCCGCAACACCG-----CGCCACCATCTCGCGGAGC 588
DB 158 ArgArgGlnProAlaAlaLeuProAlaHisArgPheArgArgHisGlyProValProAla 177
QY 589 CAATATCG-----CGACGAATACTTCAAAGTCGGTGGAGCACCGGTTTTCGCCGACT 642
DB 178 AlaTyrArgProGlyArgLeuLeuProValAlaAlaAlaGlyThrArgTyrArgArgPro 197
QY 643 GGACATCTCGCCACCGGCGTCTGCGGCAAGTATCGACGACTTCGACGGCTACTG 702
DB 198 AlaArgProThrGlyGly----- 203
QY 703 GGACAGCCATCGGCCCAACACGCGCGCATCTCCGCGGCAACATCGGCAAGG 762
DB 204 GlyArgAlaTyrArgAlaProArgThrGlyPhe-----ArgGluAlaArgThrGly 220
QY 763 TCTTCAAGCACTCGGATACACAGCAAAATCCAGACACGCGCTCTGCGCTACCGGA 822
DB 221 SerArgHisArgArgAlaGluProThrGlyArgGlyLeuGlnArgProAla-----Arg 238
QY 823 ACCCGTCGACAGTGCCTCTACCAAAAAATACAGACGGACCGCATCGACTGGCAGAG 882
```

```
DB 239 ArgThrArgIleLeuAlaGlyAlaProAla-----GlyArgGluArgLeuAlaGly 255
QY 883 CQTCCAAACCGCTGATCAG-----CGACACCCCTGCAAAAGGACTCGACCG 930
DB 256 AlaProGlyProProArgGlnProAspGlnProAlaGlnProArgLeuLeuArgArgPro 275
QY 931 CGACCGCGCAACCGCGGAGTTCGGGAGGCTCGAAGCGGCTCAACACAGCC----- 984
DB 276 ProGluProGlyAlaAlaArgCysGlnArgAlaSerArgAlaIleGlyAlaValHis 295
QY 985 CGAAAAAGCGTCTATCTGTTTACCCTA---TTTGTCTCTACAAATCCGGCACAGA 1041
DB 296 ArgGlnArgProLeuGlnGlyAspGlnArgProProArgProCysArgArgHisArg 315
QY 1042 CGCACTGGCAAACTGGTGCAGGACGATAGACGTTTACCGTCTCGCAACACTC----- 1095
DB 316 -----AlaGlyGlnHisArgHisAlaHisProTyrAlaValAlaArg 329
QY 1096 -----GCTACAGGCGACCGAGTTCGCGCGTTCATTCGCGTACGTCAA 1140
DB 330 LysArgProGlyGlyAlaProGlyArgArgGlyArgProAlaArg----- 346
QY 1141 ATACGAAACCGCTGCTCAAGCGGCATCAAACTCTACGAGCTGCAACCAACCATGC 1200
DB 347 -----ThrAlaGlyValArgArg----- 353
QY 1201 CGTCCCGCCCAAAAGACAAAGGCTGACCGCGAGCTCCGTAAACGCTGCGTCAAC 1260
DB 354 ArgProAlaHisArgArgGlnHisArgGlnHisAlaGlyAlaAspProAlaGlnArg 373
QY 1261 AACCTTCATTTGGAGCGCAAAACGATCTTCATCGGCTCATTCACCTCGACCCCGTTC 1320
DB 374 ArgGlnHis----- 376
QY 1321 CGCAGGCTCAATACCGAAATGGGCTGCTCATCGAAAGCCCAAAATCGCAGA----- 1374
DB 377 -----GlyGluTyrLeuAlaAspHisArgHisArg---ProLeuSerArgAlaCysGly 393
QY 1375 ---ACAGATGGAGCGCACCTCGCGATACCAACACACCCCAATA---CGCTACCGGTTAC 1428
DB 394 TyrProGlyGlyAlaAlaProArgArgArgHisGlyAspValHisArgGlnAlaProGly 413
QY 1429 CQTGACAAACACACCGCTGCAATGCGCATGCGATCCCGCCCGCCG 1473
DB 414 ProArgGlnProAlaProGlyGlyThrGlnArgProProHisPro 428

RESULT 29
US-09-252-991A-24941
; Sequence 24941, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 24941
; LENGTH: 475
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-24941

Alignment Scores:
Pred. No.: 4.13e-13 Length: 475
Score: 268.50 Matches: 152
Percent Similarity: 36.1% Conservative: 37
```

Best Local Similarity: 29.0% Mismatches: 166
Query Match: 9.4% Indels: 169
DB: 2 Gaps: 35

US-10-665-990A-13 (1-1561) x US-09-252-991A-24941 (1-475)

```
QY 145 CAAACCTGTCTCT-----CCTGACACATCCTCGCAAT 177
DB 28 GlnProCysProGluArgLeuProTyrLeuGlyArgProArgGlnGlyProGlyCys 47
QY 178 CGGACACCCCTCATAACACGGGCTATCCGACATCTA-----CCTGCT 222
DB 48 ProArgGlnProGlyAspArgLeuValArgArgAlaGlyGlyProGlyAlaGly 67
QY 223 CGACACCCCGACCAAGCCCTTCGCCCGCGCCCT---TATCGAATCGCGAACA 279
DB 68 ArgArgValProGlnGlySerValProProGlyHisProGlyLeuArgGlnArgAla 87
QY 280 CAGCCTCGATTGCAATACATTTGGCGCAACGACATTTCCGGCAGGCTGCTGTTCAA 339
DB 88 GluProArg-----ArgArgSerArgHisGlyGlnArgAla---Gln 102
QY 340 CCTCATGTACCTTCCGCGACAGCGCGCTCGGTACGCCCTGCTGTGTGACGACAA 399
DB 103 ProGluGlyAlaValArgArgThrGlnArgAlaGlyAlaGlyGlnProGlyThrGlnArg 122
QY 400 CACGCGGGTGGACGATCTCTGCTCGCCCTCGACGACCATCCCAATATCGAAGTGG 459
DB 123 -----GlyAspPro-----ProArgGln----- 128
QY 460 CTGTTTCAACCCCTTCGTCTACGCAAAATGGCGGCACCTCGCTACCTACCGACTTCCC 519
DB 129 ---ValGlnArgLeuArgGlnHisGlnProGluLeuPheValPro-----LeuPro 145
QY 520 CCGCCTCAACCGCGCATGACACAAATCTTTACCGCGACAAACCG----- 567
DB 146 AlaValProProValGlyProAlaArgLeuGlnArgPheProCysAlaValAla 165
QY 568 -----CGCCACCATCTCGCGGAGCGCAATATCGCGACGAATA 606
DB 166 ValArgThrValGlnProAspLeuHisLeuTyrGlnArgGlnLeuArgProAla 185
QY 607 CTTCAAAAGTCGGTGAGGACACCGTTTTCGCGACCTGGACATCCT---CGCCACCGGCG 663
DB 186 ProValArg-----ArgArgProAlaGluProGlyArgHisArgGln 199
QY 664 CGTCGTCGGCAAGTATCGACAGCTTCGACCGCTACTGGCAAGCCATCCGCGCCCAA 723
DB 200 -----ProArgLeuArgProAlaVal-----LeuGln 208
QY 724 CGCCACGCGCATCTCGACGCGGCAACATCGGCAAGGCTTTCACGACTCGGATACAA 783
DB 209 ArgHis-----ArgArgGlnTyrArgPro---AlaThrGlyAsnArgAlaGln 223
QY 784 CGACGAAACATCCAGACCGCTCTCGCTACCGGCAACCGTCGACAGCTGCGCCCT 843
DB 224 Arg-----TyrProGlyProGlyLeuPro-----AlaAlaAla 234
QY 844 CTACCAAAATATACAGACGGACCGCATCGACTGCGACAGGCTCAAAACCGCTGATCAG 903
DB 235 ArgProPheArgArgProGlyLeu-----LeuValAspGlnAspArgGlnProAspArg 252
QY 904 CGACACCCCTGCAAAAGGACTCGACCGCGCGCGCAACCGCGGATTCGCGGAGCT 963
DB 253 -----ArgValSerArgSerGlyAspLeuArgArgPro 263
QY 964 GCA-----AGACGCGCTCAAAACAGCCGAAAGCGTCTATCTGTTTCAACCTATT 1017
DB 264 AlaGlyLeuArgArgThrHisArgAlaGlnGlyArgArgLeu----- 277
QY 1018 CGTCCCTACAAATTCGGCAGACAGCAGCTGGCAAAATCTGGTCAGGACCGCATACGCT 1077
DB 278 -----TyrArgSerArgHisArgThrGlnProArgGlnGlyLysAspGlnArg 295
```

```
QY 1078 TACGCTCTGACCAA-----CTGCTTACA 1101
DB 296 ArgArgProGluProArgLeuSerPheProGlyGlnProLeuArgAlaValArgAlaArg 315
QY 1102 GGCACCGAGCTTTCGCCCGCTCCATTCCTCCGGCTA-----CGTCAAAATACCG 1146
DB 316 ProAlaArgHisLeuArgValProLeuArgLeuProAlaAlaAspArgProValPro 335
QY 1147 AAAACCCCTGCTMAACCGCGCATCAAACTCTACGAGCTGCAACCCCAACCATGCGTCCC 1206
DB 336 -----GlyGlnArgArgLeuProGlyArgArgAspArgProLeuGluAla 352
QY 1207 CGCCACAAAAGACAAAGCGCTGACCGCGCAGCTCGTAACACGAGCTGCATGCAAAACCTT 1266
DB 353 ArgArgGlnArgHisLeuGluProArgArgLeuAlaGlyHisProGlu---GlnProLeu 371
QY 1267 CATTTGGAGCGCAACGCGCATCTTCATCGCTTCATTCACCTCGACCCCGCTTCCGCAAG 1326
DB 372 His-----GlnArgLeuGlnArgLeu---ArgProArgGlnPro----- 383
QY 1327 GCTCAATACCAATGGCGT-----CGTCATCGAAAGGCC 1362
DB 384 -----ArgGlnGlyArgPheValGluProLeuGlyProGlyArgGlnLeuProPro 400
QY 1363 CAAAATCGCAGAACAGATGAGCGCACCCCTCGCGGATACACACCCCGAATACGCTACCG 1422
DB 401 GlnProArgAla-----GlyAlaAspProArgGlyGluGlu---Pro 413
QY 1423 CGTTACCTCTCGA---CAAACACACACCGCTTGCATGCGACGATCCCGCCACCGCAAAAC 1479
DB 414 ValArgProArgThrAlaValGlnGlnProAspLeu-----His 426
QY 1480 CTACCGCAACGACCGCGAAGCCAACTTTTGGAAACGCGATCGCGCAAAATCTCTATCCCT 1539
DB 427 LeuProGluArgLeuArg-----Pro 433
QY 1540 GCTGCCCATCGA 1551
DB 434 AlaLeuHisArg 437
```

RESULT 30

US-09-252-991A-24311
; Sequence 24311, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 24311
; LENGTH: 584
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-24311

Alignment Scores:
Pred. No.: 4,48e-13 Length: 584
Score: 268.50 Matches: 107
Percent Similarity: 37.6% Conservative: 59
Best Local Similarity: 24.2% Mismatches: 167
Query Match: 9.4% Indels: 109
DB: 2 Gaps: 14
US-10-665-990A-13 (1-1561) x US-09-252-991A-24311 (1-584)

```
Qy 116 CGGACGGAAGCGCTCATTTCAATCTCCAAACCTGTCCTC-----CTGACACAC 166
Db 192 ArgArgAlaAspLeuHisPheAlaValSerGlyProValPheProTrpArgSerAspAsn 211
Qy 167 ATCTGCAAAATCCGGACACACCCCTCATAAACAGCGGTATCCGACATCTACCTGCTCGAC 226
Dy 212 -----ArgPheAlaLeuAsnAspGlyValArgPhePhe----- 223
Qy 227 GACCCCAAGAGCCCTTGGCCGCGCGCCCTTATCGAATTCGCGAACAACAGCCTC 286
Dy 224 --ProArgMetLeuLeuAla-----lleGluAlaAlaGlnArgSerIle 237
Qy 287 GATTTCGAATACTACATTGGCGCAACGACATTTCCGGCAGCGTCTTCAACCTCATG 346
Dy 238 GluLeuGluLeuTrpValGluAspGlyHisCysAlaGluLeuPheLeuValAlaLeu 257
Qy 347 TACCTTGGCGGAGCGCGGTGGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 406
Dy 258 LeuAspAlaArgArgGlyValAlaValArgCysLeuPheAspGlyPheGlyCysLeu 277
Qy 407 GGGTTGGACGATCTCTGCTCGCCCTCGACAGCCATCCCAATATCGAAGTGGCGCTGTC 466
Dy 278 GlyLeuGlySerAlaTrpIleGlnArgLeuArgGluAlaGlyGlyLeuArgLeuTrp 297
Qy 467 AACCCCTTCCTACGCAATGGCGCGCACTCGC-----TACCTGACCGACTTC 517
Dy 298 AsnProLeu-----ArgTrpLysLeuThrGlyGlyAsnLeuTrp----- 310
Qy 518 CCCGCTCAACCGCGCATGCACAAACATCTTACCGCGCAACCGCGCCACCATATA 577
Dy 311 -----ArgAspHisArgLysLeuLeuValAspGlyArgLeuGlyTrp 325
Qy 578 CTGGCGGACGAATATCGGCGAGCAATCTTCAAA---GTCGGTGGAGCACACCGTTTTC 634
Dy 326 ValGlyGlyAlaGlyIleThrAspGluPheTrpGluProValSerAspValSerAlaTrp 345
Qy 635 GCGGACCTGGACATCTCGCCACCGGCGAGCTGTCGGCGAAGTATCGCAGACTTCGAC 694
Dy 346 ArgGluValMetValGluMetAspGlyProValAlaAlaAspTrpAlaAlaLeuPheGlu 365
Qy 695 CGTACTGGCAAGCAATTCGCGCCACACACGCC----- 727
Dy 366 ArgGlnTrpLeuAlaCysLeuGluGluLysAlaTrpLysProArgGluGlyMetThrLeu 385
Qy 728 ACGGGCATCATCGGCGGCAACATCGGCAAGGTCTTCAAGCACTCGGATACACACAC 787
Dy 386 ThrArgLeuProGlnProGlyAlaAlaArgGlyLeuGlyArgValAlaTrpAlaAsp 405
Qy 788 GAAACATCCAGACCGCGCTCTGCGCTACCGGNAACCGTTCGAACATCGCCCTCTAC 847
Dy 405 ----- 405
Qy 848 CAAAAATACAGACGGGAGCATCGACTGGCAGAGCGTCCAAACCGCGCTGATCAGCGAC 907
Dy 406 -----AlaArgGlnHisArgAspIleLeuGln 414
Qy 908 ACCCTCGAAAGGACTCGACGCGACCGCGCAACCGCGATTCGCGGAGGCTGCAA 967
Dy 415 SerLeuValArgAlaLeuAsnGlySerArgArg----- 426
Qy 968 GACGGCTCAACACGCGCAAAAGCGTCTATCTGGTTTCAACCGTATTCGTCCTTACA 1027
Dy 427 -----lleTrpLeuAlaTrpProTrpPheLeuProThr 437
Qy 1028 AAATCCGGCACAGCGCACTGGCAAACTGGTGGAGGAGCGGATAGAGCTTACCGTCTG 1087
Dy 438 TrpLysValArgArgAlaLeuArgLysAlaAlaGlnArgGlyValGluValArgLeuLeu 457
Qy 1088 ACCAACTCGCTACAGCGACCGAGCTTGGCGCGCTTCATTCGGGTACGTCAAATACCGA 1147
Dy 458 LeuAlaGly---ArgLeuThrAspHisAlaProValArgTrpAlaGlyGlnArgTrpTrp 476
Qy 1148 AAACCGCTCTCAAGCGCGCATCAAACTCTACGAGCTGCAACCAACCAACATCGCGTCCC 1207
```

```
Dy 477 ProArgLeuLeuArgAlaGlyValArgIleHisGluTrpGlnProArgPhe----- 493
Qy 1208 GCCACAAAGACAAAGCGCTGACCGGAGCTCCGTAAACACGCTGCATGCCAAACCTTC 1267
Dy 494 -----LeuHisLeuLysMetVal 499
Qy 1268 ATTGTGACGGCAACGCAATCTTCATCGGCTCATTTCAACCTCGACCCCGTTCGCACGG 1327
Dy 500 MetValAsp---AspTrpValSerValGlySerCysAsnPheAspHisTrpAsnLeuArg 518
Qy 1328 CTCAATACCGAAATGGCGTCTCATCGAAAGCCCAAAATCGCAGAACACATCGAGCGC 1387
Dy 519 PheAsnLeuAspAlaAsnLeuGluAlaLeuAspProAspPheThrAsnGluAlaAla 538
Qy 1388 ACCCTC 1393
Dy 539 SerLeu 540

RESULT 31
US-09-252-991A-17125
; Sequence 17125, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 17125
; LENGTH: 579
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-17125

Alignment Scores:
Pred. No.: 4,89e-13 Length: 579
Score: 268.00 Matches: 111
Percent Similarity: 37.7% Conservative: 60
Best Local Similarity: 24.5% Mismatches: 186
Query Match: 9.4% Indels: 96
DB: 2 Gaps: 13

US-10-665-990A-13 (1-1561) x US-09-252-991A-17125 (1-579)
Qy 95 TGGTTGCCCCCACTGGAAGAACGACGAAAGCGCTCATTTCAATCTTCCAAACCTGTC 154
Dy 186 TrpArgProTrpValGluAlaLeuAlaArgGlnValSerGlyTrpLysGly--- 204
Qy 155 CTCTGGAACAACATCTCGCAATCGGCACACCCCT-----CATAACAACGGGTATCC 208
Dy 205 ---LeuLysAlaLeuValArgMetThrArgThrProThrLeuAlaAsnAsn--- 220
Qy 209 GACATCTACCTGCTCGACGACCCCGACGAAACCGCTTGGCGCGCGCGCTTATCGAA 268
Dy 221 ArgValArgLeuLeuValAsnGlyGluAlaSerPheGluAlaMetPheLysAlaIleSer 240
Qy 269 TCTCCGCAACACAGCCTCGATTTCGAATCTACATTTGGCGCAACAGCATTTCCGGCAGG 328
Dy 241 AlaAlaArgGlnValIleLeuValGlnPhePheIleValArgAspAlaLeuGlyGln 260
Qy 329 CTGCTGTTCAACCTCATCTGCTTGGCGCAACCGCGCGTGGCGGTACGCTGCTGTTG 388
Dy 261 ArgLeuGlnGlnLeuLeuGluArgAlaAlaAsnGlyValGluValPhePheLeuTrp 280
Qy 389 GACGACACAAACACGCGCGGTGGACGATCTCTGCTCGCGCTCGACAGCCATCCCAAT 448
Dy 389 GACGACACAAACACGCGCGGTGGACGATCTCTGCTCGCGCTCGACAGCCATCCCAAT 448
```


Db 281 AspAlaIleGlySerHisAlaLeuProHisArgTyrValGluArgLeuArgGlnGlyGly 300
Qy 449 ATCGAAGTGGCGCTCTCAACCCCTCTGCTCTACGCAAAATGGCGGCACCTGGCTACCTG 508
Db : : : : :
Db 301 ValGlnMetHisGlyPheSer-----ThrGlySerGlyMetLeu 313
Qy 509 ACCGACTTCCCGCGCTCAACCGCGCATGACAAATCTTTACCGCGCAGCAACCGC 568
Db : : : : :
Db 314 AsnArgPhe--GlnValAsnPheArgAsnHisArgLysValValValAspGlyGlu 332
Qy 569 GCCACCATACTCGGCGAGCAATATCGCGACGAACTACTTCAAGTCTGGTGGAGACACC 628
Db : : : : :
Db 333 CysGlyPheValGlyGlyHisAsnValGlyValGluTyr-----LeuGlyGluLysPro 350
Qy 629 GTTTTCGCC-----GACGTGCATCTCGCGCACCGCGAGCGTCTCGCGGGAAGTA 679
Db : : : : :
Db 351 ProLeuAlaProTyrArgAspThrHisMetGluLeuArgGlyProAlaValAlaCysLeu 370
Qy 680 TGCACGACTTC-----GACCGTACTGGCAAGCCATTCGCCCCACACGCCACGCGC 733
Db : : : : :
Db 371 GlnGluSerPheAlaGluAspTyrTyrAlaThrHisSerLeu----- 385
Qy 734 ATCATCCGCGGCAACATCGGCAAGGTCTTCAAGCACTCGGATACAGCAAGAAC 793
Db : : : : :
Qy 794 TCCAGACGCGCTCTCGCTACCGCGAAGCGTCGAACAGTCGCCCTCTACCAAAA 853
Db : : : : :
Db 386 -----ProProLeuIleLeuProGlnTyrAspSerGluGlyAlaLeuCysGlnVal 403
Qy 854 ATACAGCGGACGATCGATTCGAGAGCGCTCCAAACCCCGCTGATCAGCGACACCCCT 913
Db : : : : :
Db 404 ValAlaSerGlyProAlaAspAlaGlnGluThrCysSerLeuPhe----- 418
Qy 914 GCAAAAGACTCGACCGCGCGCGCGCAACCGCGATTCGCGGAGGCTCAAGACGCG 973
Db : : : : :
Db 419 -----PheValGluMet 422
Qy 974 CTAAACAGCCCGAAAAAGCTCTATCTGTTTCACCCCTATTCGCTCCCTACAAAATCC 1033
Db : : : : :
Db 423 IleAsnAlaAlaHisGluArgValTyrPileThrSerProTyrPheValProAspGluAla 442
Qy 1034 GGCACAGCGACTGGCAAACTGGTGAGGACGCGATAGAGCTTACCGTCTCTGACCAAC 1093
Db : : : : :
Db 443 ValMetAlaAlaLeuArgLeuAlaValLeuArgGlyValAspValArgLeuLeuIlePro 462
Qy 1094 TCGCTACAGGCGACGAGTTCGCGCGCTCCATTCGCGCTACGTCAATACCGAAACCG 1153
Db : : : : :
Db 463 Ser-----ArgProAspHisArgThrValTyrAlaAlaSerSerLeuTyrAlaLeuGlu 480
Qy 1154 CTGCTCAAAGCGCGCATCAAACTCTACGAGCTGCAACCCCAACCATGCGCTCCCGCCACA 1213
Db : : : : :
Db 481 AlaIleArgAlaGlyValLysValPheArgTyrGlnProGlyPhe----- 495
Qy 1214 AAAGCAAAAGCGCTGACCGGCGAGCTCCGTAACCGACTGCAATGCCAAACCTTCATTGTG 1273
Db : : : : :
Db 496 -----LeuHisGlnLysValValLeuVal 503
Qy 1274 GACGGCAACCATCTTCATCGGCTCATTCACCTGATTCACCTCGACCCCGCTTCGCGACGCTCAAT 1333
Db : : : : :
Db 504 AspArgAspThrAlaAlaValGlySerAlaAsnLeuAspAsnArgSerPheArgLeuAsn 523
Qy 1334 ACCGAAATGGCGTCGTCATCGAAAGCCCAAAATCGCAGAAACAGATGGAGCGCACCTC 1393
Db : : : : :
Db 524 PheGluValMetValValThrValAspGlyPheAlaGlyGluValGluAlaMetLeu 543
Qy 1394 -----GCCGATACCAACCCCGAA 1411
Db : : : : :
Db 544 GluAlaAspPheAlaGluSerLeuGluPheThrProGlu 556

RESULT 32

US-09-252-991A-30799

; Sequence 30799, Application US/09252991A

; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252.991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 30799
; LENGTH: 997
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-30799
Alignment Scores:
Pred. No.: 6,03e-13 Length: 997
Score: 268.00 Matches: 156
Percent Similarity: 37.0% Conservative: 42
Best Local Similarity: 29.2% Mismatches: 175
Query Match: 9.4% Indels: 162
DB: 2 Gaps: 34

US-10-665-990A-13 (1-1561) x US-09-252-991A-30799 (1-997)

Qy 163 CAACATCTGCAATCCGGCACACCCCTCATACACAGCGGTATCCGACATCTACCTCT 222
Db : : : : :
Db 215 GlnHisProAlaGly--AlaArgProGluArgPheArgProValArgArgAlaProAla 233
Qy 223 CGACGACCCCAACAGCCCTTTCGCGCGCGCGCCCTTATCGAATCTGCGGACACAG 282
Db : : : : :
Db 234 ArgArg-----LeuArgProAlaGln--LeuLeuLeuCysArgGlnArg 247
Qy 283 CCTCGATTGCAATACTACTATTGGCGCAACGACATTTC-----CGGACGCTCT 333
Db : : : : :
Db 248 -----ArgGlnArgHisLeuProLeuArgArgGlnArgPro 259
Qy 334 GTTCAACCTCATGTACTTCGCGCAGAACCGCGGTGCGGTACGCTGTGTTGGACGA 393
Db : : : : :
Db 260 ProAspProLeuHisPro--ArgArgGlyArgGlnArgProAlaProAlaArgHisArg 278
Qy 394 CAACAAACAGCGCGGGTTGGACGA-----TCTCTGCTGCGCCCTCGACAGCCA 441
Db : : : : :
Db 279 ArgArgHisArgArg--GlyHisLeuArgHisProGlnProGlyGlyGlnArgGlnArg 297
Qy 442 TCCCAATATCGAAGTGGCGCTTCAACCCCTTCGCTACGCAAAATGGCGCGACTCGG 501
Db : : : : :
Db 298 ArgProGlyArgGlnArgProAlaIleProAlaArgLeuLeuLeuPheAlaThrAla 317
Qy 502 CTACTGACCGACTTCCCGCGCTCAACCGCGCATGCAACAAATCTCTTTACCGCGCA 561
Db : : : : :
Db 318 ArgProGln-----ProAlaThrHisArgAlaGlnArgArgLeuTrpArg--- 332
Qy 562 CAACCGCGCCACCATCTACGCGGACGCA-----TATCGG 597
Db : : : : :
Db 333 ---ProArgHisProGlyArgArgValProValGluProAlaHisArgProGlyTyrArg 351
Qy 598 CGAGCAATCTTCAAGTCGCTGAGGACACCGTTCGCGACCTGCGACCTGCTCGCCAC 657
Db : : : : :
Db 352 GlnProAspLeuArgArgAlaGlyHisArgTyrArgLeuLeuArgHisProArg--- 370
Qy 658 CGGACGCTCGTGGCGAAGTATCGCACGACTTCGACCGCTACTGCGCAAGCCATTCGCG 717
Db : : : : :
Db 371 -----ProArgGlnProHisAlaProGlyGlnProLeuLeuGlnProProAspArg 388
Qy 718 CCACACGCCACCGCATCATCCGACGCG-----CNACATCGCGCANGGCTCTCA 768
Db : : : : :
Db 389 -----GlnProGlnGlnLeuLeuProGlyAlaGluGlnGlySerArg 402


```
Db 232 -----ArgLeuGluArgAsnLeuGlu-----ProAla-----Gln 242
Qy 823 AACGTCGAACAGTCGCCCTCTACCAAAAATACAGACGGACGCATCGACTGCAGAG 882
Db 243 GluArgArgAspLeuProAlaValAlaAlaHisProArgGly-----AlaGln 258
Qy 883 CGTCAAAACCGCTGTAT-----CCGCGACGGCGGAAACCGCGGATTCGGCGAGGCTGCA 915
Db 259 ArgProGlyProAlaAspProLeuArgArgGlyValLeuArgProGluGlnHisGlnAla 278
Qy 916 AAAAGGACTCGA-----CCGCGACGGCGGAAACCGCGGATTCGGCGAGGCTGCA 966
Db 279 LeuGlyGluArgAlaGlyLeuProArgProProArgLeuAlaAsp---ArgProAlaGln 297
Qy 967 AGACGCGTCAACACGCGGAAAGGCTATCTGCTGTTTCAACCTATTCCTCCCTAC 1026
Db 298 ProArgAlaAlaAlaArgAlaHisArgAlaGlyAlaGly----- 310
Qy 1027 AAAATCCGCGACAGACGCTGGCAAACTGGTCGAGCAGCGCATAGACGTTACCGTCT 1086
Db 311 -----GluArgGlnGlyProAspGlyGlyArgGlyProAlaAlaAspArgPro 326
Qy 1087 GACCAACTCGTACAGGCGAGCGAGTTGCGCGCTGCATTCGGGCTACGTCAAATACG 1146
Db 327 ArgProLeuGlnAlaHisGlnArgGln---ProArgProTyrAsnTyrArgHisAlaPro 345
Qy 1147 AAAACCGTGTCAAAGCGGCATCAAACTCTACGAGCTGCAACCCACCATGCGTCCC 1206
Db 346 GlnGlyGlnGln-----AlaProAlaThrProAlaArg--- 357
Qy 1207 GCCCAAAAGACAAAGGCTGACCGCAGCTCGTACACGCTGCATGCAAAACCTT 1266
Db 358 -----ArgAlaLeuProAla-----Val 363
Qy 1267 CATTGTGACGGCAAAACGATCTTTCATCGCTCATTTCAACCTCGACCCCGCTCCGACG 1326
Db 364 AlaThrGlyArgArg----- 368
Qy 1327 GCTCAATACCGAATGGCGCTGTCATCGAAGCCCAAAATCGCAACAGAT---GGA 1383
Db 369 ---ArgIleArgHisProArgArgGluArgProGlySerArgAlaProLeuAla 387
Qy 1384 CGCACCTCGCGATACACACCGCATACGCTACCGCTTACCGCTTCGACAAACACAA 1443
Db 388 AlaHisProArgArgPheGlnArgAlaValArgHisProLeu----- 401
Qy 1444 CGGCTGCAATGGCAGCATCCGCGACCCGCAAAACCTTACC-----GAACGAAC 1494
Db 402 ---ProAlaAspLeuTyrGlnArgLysProArgCysGlnProLeuProGlyArgGln 420
Qy 1495 CGAAGCCAACTTTGGAACCGCATCGCGCAAAATCCTATCCCTGCTGCCCATCGAAG 1554
Db 421 ArgArgArgProProAspAlaThrArgArgArgAlaValProGlyGlnGlyArg 440

RESULT 34
; US-09-252-991A-17231
; Sequence 17231, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252.991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 17231
; LENGTH: 2294
; TYPE: PR
```

```
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-17231
Alignment Scores:
Pred. No.: 1,1e-12 Length: 2294
Score: 266.50 Matches: 185
Percent Similarity: 30.6% Conservative: 43
Best Local Similarity: 24.9% Mismatches: 223
Query Match: 9.3% Indels: 293
DB: 2 Gaps: 39

US-10-665-990A-13 (1-1561) x US-09-252-991A-17231 (1-2294)
Qy 121 GGAAGCCGTCATTTCAATATCTTCCAAACCTGCTCCTCGACAAACATCTCGCAATCCG 180
Db 86 GlyProAlaArgTyrHisArgLeuProArgProAlaAlaGluAlaGlyArgHisPro 105
Qy 181 GCACACCCCTCAAAACACGCGCTATCCGACATCTACCTGCTCAGCACCC- 231
Db 106 ArgHisGlyArgGlyArgGlyProArgArgProLeuArgGlnProGlyAlaGly 125
Qy 232 -----CCAGAACCCCTTGCCTGCGCCG- 252
Db 126 SerAlaArgArgGlnCysArgLeuArgGluProArgProGlyGlyProGlyGlnProGly 145
Qy 253 -----CGCGCCCTTATCGAATCTCCGA- 276
Db 146 ValProAlaAspArgAlaAlaArgProArgArgIleArgArgGlyGlyArgAla 165
Qy 277 -----ACACAGCTCGATTTGCAATACTACATTTGCGC 309
Db 166 GlyProArgGluAlaArgGlyProGlyProGlyProGlyArgGlyLeuProAlaAlaGly 185
Qy 310 CAACGACAT-----TTCCGCGAGGCTGCTGTTCACCTCAT 345
Db 186 AlaArgHisCysArgArgAlaAlaProGlyValLeuArgGlnGlyAlaAlaGlyProAla 205
Qy 346 GTACCTTCCGCGAGAACG---CGCGCTGCG- 372
Db 206 ProProAlaArgArgSerProArgAlaAlaThrGlyGlyGlyGlyGlyGlyArgLeu 225
Qy 373 -----CGTACGCTGCTGTGGACGACAAACACACGCGCG- 408
Db 226 GlyAsnArgArgLeuProAlaAlaAlaThrGlyGlnProArgArgGlyProHisTrpSer 245
Qy 409 ---GTTGGACGATCTCTGCTCGCCCT- 432
Db 246 ProValProArgAlaProAlaAlaProAlaGlyAlaValProArgThrGlyGlyProGly 265
Qy 433 -----CGACAGCCATCCCAATATCGAAGTGCCT---GTTCAA 468
Db 266 ArgArgValLeuTyrLeuHisArgGlnProAlaAlaSerArgGlyValProGlyValGln 285
Qy 469 CCCTTCTGCTCAGCAAAATGCGCGC- 495
Db 286 ProArgArgProGlyArgGlnLeuArgProGlyAspAlaArgProAlaProGlyAlaGly 305
Qy 496 -----ACTCGCTACTGACCGA- 513
Db 306 AspGlyAspArgArgGlnGlyAspArgHisProAspArgArgGluGlyLeuArgAla 325
Qy 514 -----CTTCCCGCTCAACCGCCCGCATGTCACAAACATCCTT-----TAC 555
Db 326 ArgArgAspAlaGlnArgProGlnProAspArgGlnArgGlnArgProGlyGlyAla 345
Qy 556 CGCGCAACACCGCGCCACCATCTCGCGCGACGCAATATCGCGACGAATCTTCAAGT 615
Db 346 ArgArgArgProArgArgGluProAlaAlaArgAlaLeuProAlaAlaGly-----GlnGly 363
Qy 616 CGGTGAGGACACCGTTTCCCGACCTGGACATCT- 651
Db 364 ArgLeuGlyGlyAlaHisProArgProGlyGlnProAlaThrAlaGlyTyrArgProAla 383
```

```
QY 651 ----- 651
Db 384 AlaAspAlaGlyGlyAlaAlaAspProLeuProArgArgLeuHisLeuPheArgThr 403
QY 652 -----CGCCACCGGAGCGTGGT 669
Db 404 GlyProArgGlnArgAlaLeuGluAlaAlaAspAlaLeuArgArgLeuArgLeuProHis 423
QY 670 CGCGGAAGTATCGCAGCATTCGACCG---CTACTGGGC----- 705
Db 424 ArgArgProValProArgAlaGluProGlyLeuLeuGlyHisSerArgIleSerAsnAsp 443
QY 706 -----AAGCATTCCGC 717
Db 444 AlaSerHisArgArgSerLeuArgHisProGlyArgArgGlnArgGlyProAlaArg 463
QY 718 CCACAACGCCAGCGCATCATCGCAGCGGCAACATCGCAAGGGTCTTCAAGCACTCGG 777
Db 464 ArgProHisHisAspHisAlaAlaSerArgArgProArgProGlyThrGlyArgArgPro 483
QY 778 ATACAACGACGAAACATCCAGACACGCGCTCTCGCTACCG-----CGAAAC 825
Db 484 ProArgArgArgSerLeuHisAlaAlaAlaAspGlyGlnProGlyAlaAlaGluArgPro 503
QY 826 CTCGGAACAGTCGCCCTCTACAAATAATACAGACGGGACGCATCGACTGGCAGAGCGT 885
Db 504 ArgArg-----ArgProGlnProAlaGlyAla--GlyArgArgProThrAlaGlyAsp 520
QY 886 CCAACACCGCTGATCAGCGACAC-----CCCTGCAAAAGGACT 924
Db 521 AlaHisProAlaGlnHisHisArgProSerGlyAlaGlyGlnProAlaAlaGluAlaThr 540
QY 925 CGA-----CCGCGACCGCGCAA-----ACCGCGATTGCGGAGGCTCGAAGA 969
Db 541 AlaArgLeuProAlaProValArgGlyGluGlyThrProGlyArgArgGlyAlaGlyArg 560
QY 970 CGCGCTCAACAGCCGGAAGAGCGTCTATCTGTTTCACCTATTGCTCCCTACAAA 1029
Db 561 SerAspAlaGlyAlaLeuArg-----AlaLeuHisArgPro----- 572
QY 1030 ATCCGCGACAGCGACTCGCAAAATCGTGCA---GGACGCGCATAGAGTTTACCGTCT 1086
Db 573 -----ArgArgSerGlyAlaGlnHisAlaValAlaGlyGlnHisGlnArgLeu----- 587
QY 1087 GACCAACTCGTACAGGCGACGAGCGTTCCCGCGTCCATTCCGCTACGTCAAATACCG 1146
Db 588 GlyGlnAlaGluProAlaAspHisProAlaGlnArgGlyLeuGlyArgArgGluGlyLeu 607
QY 1147 AAAACCGTGTCAAAGCCGCGCATCAACTCTACGAGCTGCAACCAACCATGC----- 1200
Db 608 ProValAla-----ArgThrLeuProAlaGluProAlaProAlaProAla 622
QY 1201 -----CGTCCCGGCCACAAAGACAAAGGCGCTGACCGG 1233
Db 623 ProAlaGlyAlaAlaValProValHisGlnProArgLeuArgArgProLeuProGlyHis 642
QY 1234 CAGCTCCGTAACAGCCT-----GCATGCCAAACCTTTCATTGTGGACGGCAACG 1284
Db 643 AlaArgArgProGlnProThrGlyGlyAlaAlaArgAlaHisArgGlyGlyAspProGln 662
QY 1285 CATCTTCATCGG-----CTCATTCACCT----- 1308
Db 663 HisProArgArgValArgAlaArgThrValAlaAlaLeuAlaArgProGlnArgAspAla 682
QY 1309 CGACCCCGCTTC-----CGCAGCGCTCAATACCGAAAT----- 1341
Db 683 ArgProProPheAlaIleProAlaAlaValAlaGlyArgTyArgHisArgProGlyLeuAla 702
QY 1342 -----GGCGTCGTCATCGA----- 1356
Db 703 AlaGlyLeuAlaValArgProAlaAspGluAlaGlyArgArgArgAlaGlyValGln 722
QY 1357 ---AAGCCCCAAATCGCAGAACAGATGGAGCGCACCTCGC----- 1395
```

```
Db 723 AlaHisProArgHisArgArgAspProGlyAlaGlyHisArgProGlyGlyAlaAla 742
QY 1396 -----CGATACCACACCGAATACGCTTACCGCTTACCTCGACAAACACACCG 1446
Db 743 AlaGlyAspArgAlaAlaThrProGlyAlaLeuProArg-----ArgArgHisGlnGly 760
QY 1447 CCTGCA-----ATGGCACGATCCCGC-----CACCCGAAAAACCTA 1482
Db 761 ProAlaGlyGlyGlyGlyArgArgArgProLeuGlyGlyAspHisProArgArg 780
QY 1483 CCGACGAAACCCGACGCAAACTTGGAAACGATCGCGCAAAAATCCTATCCCTGCT 1542
Db 781 AlaValArgGlnArgGlnArg-----GlnArgAlaArgValProAlaAlaAla 798
QY 1543 GCCCATCGAAGG 1554
Db 799 AlaHisArgArg 802

RESULT 35
US-09-252-991A-27864
; Sequence 27864, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 27864
; LENGTH: 777
; TYPE: PR
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-27864

Alignment Scores:
Pred. No.: 8,686-13 Length: 777
Score: 265.50 Matches: 158
Percent Similarity: 33.9% Conservative: 49
Best Local Similarity: 25.9% Mismatches: 202
Query Match: 9.3% Indels: 202
DB: 2 Gaps: 31

US-10-665-990A-13 (1-1561) x US-09-252-991A-27864 (1-777)
QY 120 CGGAAGCGCTCATTTCAATA-----CTTCCAACCTGTCC 155
Db 197 ArgLysSerValSerAlaIleHisSerArgGluLeuAlaAspProGluProArgAlaPro 216
QY 156 TCCTGGACAAACATCTCGCAATCCGGACACCCCTCATA---ACAACGGGCTAT---CCG 209
Db 217 ThrLeuAlaSerSerValArgGlyGlyThrAlaLeuIleArgLeuSerGlyTrpArgPro 236
QY 210 ACATCTACTGCTCGACGACCCCGACGAG----- 239
Db 237 GlnSerSerValAlaAlaValGlyThrArgAlaSerThrArgAlaGlyGlyAlaThrAla 256
QY 240 ---CCCTTGGCGCGCGCGCGCTTATCGAATCTGCGG---AACACAGCCTCGATTTGC 293
Db 257 TrpProProTrpAlaSerLysAlaSerGlnAlaProSerAsnArgGlnAlaSerCys 276
QY 294 AATACTACATTTGGCGCAACGACATTTCCGGCAGCGCTCTGTTC----- 338
Db 277 AsnAspArgGlnGly-----AsnAlaGluAlaGlyCysAlaGluAlaIleThrAspPro 294
QY 339 -----ACCTCATGTACCTTGGCGCAGAACCGCGGTGCGGTAC 377
```

Db 295 LeuileValMetArgAlaProSerProTyrThrLeuProArgProSerAlaProPheVal 314
QY 378 GCCTGCTGTGGACACAAACACGCGGGTGTGGACGATCTCTCTGCTCGCCCTCGACA 437
Db 315 GlyLeuAlaGlyLysThrProThrSerSerArgTrp-ArgArgLeuSerArgGluCysAr 334
QY 438 GCATCCCAATATCGAAGTGGCCCT- - - - - GTTCAACCCCTT- - - - - CGTCTTAC 482
Db 334 ggluThrGlyLysGluValSerProGlyCysValAsnLeuLeuAspTrpAsnThrProG1 354
QY 483 GCAATGGCGGCACCTCGCTACCTGACGAGTCTCCCGCCCTCAACG- - - - - 531
Db 354 uProProThrArgHisAlaThrSerLeuProAlaProProProArgArgArgLe 374
QY 532 -CCGCATGCACAAACAACTTTTACCGCGCAACCGCGCACCATCTCGCGCGACGCA 590
Db 374 uProAspAlaArgArgTrpProGlyGlnArgArgProArgHisArgGlnArgThr 394
QY 591 ATATCGCGCAGCAATA- - - - - CTTCAAAAGTCGGTGAGGACAC- - - - - CGTTT 632
Db 394 r---ArgHisArgLeuGlyProGlyProGlyLeuArgSerProProGlyArgProArgTh 413
QY 633 TCGCCGACCTGACAT- - - - - CCTCGCACCGGAGCGTCTGTCGGCAAGTATC- - - - - GC 683
Db 413 rLeuAlaAlaGlyHisArgProProTrpArgGlnArgLeuArgThrGlyThrHisProG1 433
QY 684 ACGACTTCGACCG- - - - - CTACTGGGCAAGCCATTCCG 716
Db 433 yArgLeuArgProGlyGlyAspArgGlyArgLeuArgArgAlaGlyProGlyAspAs 453
QY 717 CCCACACGCCACGCGCATCATCTCCGACGCGCAACATCGGCAAGGCTTCAAGCACTCG 776
Db 453 pProArgArgGlnAlaGlyArgProProArg- - - - - 463
QY 777 GATACAACGCAAAACATCCAGACACGCGCTCTCTCGCTACCGCGAAACCGTCGAACAGT 836
Db 464 ---GlnArgAlaGlyProAspHisArgCysArgProAlaProGlyLleArgArgProG1 482
QY 837 CGCCCTCTACCAAAATACAGAGCGGACGACATCGACTGCGAGCGTCCAAACCCGCC 896
Db 482 nAlaHisAlaGluSerArgArgSerGlyThrHisArgLeuValGlnArg- - - - - 498
QY 897 TGATCAGCGACACCCCTGCAAAAGGACTTCGACCGCGACCGCGGCAACCGCGATTGCGG 956
Db 499 - - - - - ArgLeuHisProArgArgThrGluAsp- - - - - 507
QY 957 GGAGCTCGACAGCGCTCAACACAGCCCGAAAGCGTCTATCTGTGTTTCACCCCTATT 1016
Db 508 -ProAlaArgHisArgAlaHisPro- - - - - 515
QY 1017 TCGTCCCTACAAATCCGGCAGCAGACGCACTGGCAAACTGGTCAGGACGCGATAGACG 1076
Db 516 - - - - - HisAspProProGlyGlnCysAlaProGlyArgHisLeuAr 529
QY 1077 T- - - - - TACCGTCTCT- - - - - GACCAACTCGGTAC 1100
Db 529 gAspProHisProAlaGlyAspHisArgProGlyGluLysProAlaAspGlnProAlaTh 549
QY 1101 AGCGACCGCA- - - - - CGTTGCGCGGTCTCCATTCCGGCTA- - - - - 1134
Db 549 rHisHisArgProValProArgAspGlnAlaArgHisProLeuProAlaProArgProG1 569
QY 1135 - - - - - CGTCAATACCGAAACCGTGTCTCA 1160
Db 569 yHisGlyThrProThrGlyGluAspProAlaProGlnArgLeuProArgSerAlaArgAl 589
QY 1161 AAGCGCGCATCAAACTCTACGAGCTGCAACCAACCATCGCTCCCGCCCAAAAGACA 1220
Db 589 aglyLleHisProValLeuArgGlyGluInProGlyThrGlnAlaProHisArgHi 609
QY 1221 AAGGCTGACCGGACGCTCCGTAACACGAGCTGCATGCCAAACCTTCATTGTGACGCGCA 1280
Db 609 sProProGlyProAlaLeuArgGlnArgProAlaLeuArgProAlaGlyCysArgArgG1 629

QY 1281 AACGCATCTTCATCGG- - - - - 1296
Db 629 nProAspLeuArgArgAspGlyHisArgGlnGlyProAlaProGlyGlyAlaLeuArgLe 649
QY 1297 - - - - - -CTCATTCAACCTCGACC 1313
Db 649 uArgGlyArgSerGlyGlnGluLeuArgAspSerProArgCysGlnArgGlnProArgPr 669
QY 1314 CCCGTTCCGACGGCTCAATACCGAAATGGCGTCTCATCGAAAGCCCAAAATCGCAG 1373
Db 669 o- - - - - AlaAspProLeuArgArgArgProArg- - - - - ArgAr 681
QY 1374 AACAGATGGAGCGCACCTCGCGATACACACCCGAAATACGCCTACCGCT- - - - - 1425
Db 681 gAlaGluGlyAla- - - - - SerLeuHisLeuProArgGlyLysG1 694
QY 1426 - - - - - TACCCTGCACAAACACACACCGCCTGCAATGCACGA- - - - - TC 1463
Db 694 nLeuProSerGlyGlyValProGlnArgArgGlnProAlaValAlaArgArgProG1 714
QY 1464 CGGCCACCCGAAACCTACCGGACGAAACCCGAGCAAACTTTGAAACGCGATCGCG 1523
Db 714 yArgArgAspGlnGlyLeuProArgArg- - - - - ArgHisArgAr 727
QY 1524 CAAAAATCTATCCTCTGCTGCCCATCGAAGG 1554
Db 727 gProValGlnArgProAlaGlyArgArgArg 737

RESULT 36

US-09-252-991A-27514
; Sequence 27514, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 27514
; LENGTH: 627
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-27514

Alignment Scores:
Pred. No.: 9.6e-13 Length: 627
Score: 264.50 Matches: 154
Percent Similarity: 37.0% Conservative: 36
Best Local Similarity: 30.0% Mismatches: 211
Query Match: 9.3% Indels: 113
DB: 2 Gaps: 29

US-10-665-990A-13 (1-1561) x US-09-252-991A-27514 (1-627)

QY 112 AGAAGCAGCGAAAGCGCGTCATTTCATCTTCCAA- - - ACCTGCTCTCTGGACAACAT 168
Db 173 ArgThrAspGlyArgAspArgValArgHisArgLeuArgAlaProProGlyGluGln 192
QY 169 CCTGCAATCCGCGACACACCCCTCATACACACGGGCTATCCGACATCTACCTGCTCGACGA 228
Db 193 Ala- - - - - GluGlyHisProGly- - - - - ProGlyArgGlu 202
QY 229 CCCCCAGGAAGCCCTTCGCGCGCGCGCCCTTATCGAATCTGCCGAAACACAGCCTCGA 288
Db 203 GlyProGlyHisAlaArgArgSerArgAlaProAlaAlaGlyArgArgArg- - - - - 220

```
QY 289 TTTGCAATACTATATTGGCGAAGACATTTTCGGCAGGCTGCTTTCACCTCATGTA 348
Db 221 -----HisLeu---GlnArgArgA-gArgAla-----Leu 230
QY 349 CCTTGGCGAAGACGGCGGTGCG---CGTAGCGCTTCTGTGTGGACGACAAACACGCG 405
Db 231 Pro-----ArgIleArgAlaAspArgHisPro-----GlyAspProAspProAsp 246
QY 406 CGGTTGGACGATCTCTCTGCTCGCCCTCGACGCCATCCCATAT---CGAAGTGGCGCT 462
Db 247 GlyLeuGlyArgAsnProArgArgProProAlaAspGlyArgHisGlyArgProAlaAsp 266
QY 463 GTTCAACCCCTTCGCTCCTACGCAATGGCG-----CGCACTCGGCTACCTACCGACTT 516
Db 267 LeuAlaProLeuArgGlnArgHisProAlaGlyIleArgProGlyAlaArgHisArgGln 286
QY 517 CCCCCGCTCAACCGCGCATGACAAACAATC---CTTTACCGCGCACACCGGCCAC 573
Db 287 ProLeuGlyGlnProProHisArgGlnArgA-gGlyLeuTyArgArgSerProLeuHis 306
QY 574 CATACTCGG---CGGACGCAATATCGGCGAAGAACTTTCAAAAGTCGGTGAGCACCGCT 630
Db 307 ProCysArgHisArgThrAspProAspArgProGlyValHisAlaArgProGlyHisArg 326
QY 631 TTTGCGGACCTGGACATCTCGCCAC-----CGGACGCTGCTCGG 672
Db 327 LeuArgArgArgArgProArgArgIleProArgSerGlyProArgMetGluGlyArg 346
QY 673 CGAAGTATCGCAGCTTCGACCGCTTACGCGCAAGCCATTCGCGCCACCAAGCCAC--- 729
Db 347 ArgGlnAlaGluValAlaArgArgLeuAlaGlyLysLeuProAlaAlaGlnThrHisAla 366
QY 730 ----GGCATCATCGGCGGCAACATCGGCAAGGTCTTCAAGCACTCGGTATACAAGA 786
Db 367 AlaAlaGlnAspProLeuArgGln---ArgAlaGlyGluThrAlaAlaArgLeuArgArg 385
QY 787 CGAACAATCCAGACGCGCTCTGCGCTACCGCA----- 822
Db 386 AspGluArgValLeuArgGlnGlyCysLeuLeuArgGlnHisArgProValAlaAsp 405
QY 823 AACCGTTCGAACAGTGCGCCCT---CTACCAAAAAATACAGCGGACGCATCGACTGGCA 879
Db 406 ArgArgArgThrValProProCysLeuPro-----ThrAlaProLeuAsp 420
QY 880 GAG---CGTCCAAACCCGCTGATCAGCGACACCCCTGCAGAAAGGACT----- 924
Db 421 GlnLeuArgProGlyArgProAlaGlyLeuAspHisSerArgAlaGlyGlyGln 440
QY 925 ---CGACCGGACCGCGCAACCCCGGATTCGCGGAG----- 960
Db 441 GlyArgProGluProProGlyGlyGlyAlaLeuArgArgLeuArgLeuProValHisAsp 460
QY 961 -----GCTGCAAGACGCGCT---CAACAGCCCGAAAAAGGCTATCTGTTTC 1008
Db 461 ArgGlyThrGlyGlyArgArgAlaValGlnAlaLeuTyProArgThrGlyGln 480
QY 1009 ACCCTATTGCTCC-----TACAAATCCGCGACAGCGACTGCGCAA 1053
Db 481 LeuLeuProGlyProAspProProGlyAlaAlaArgLeuArgHisGlyLeuLeuArgAla 500
QY 1054 ACTGTTGACGAGCGGATAGAGTTTACCGTCTCGACCACTCGCTACAGGCGACGAGT 1113
Db 501 AlaGlyLeuArgGlnHisGlnArgProArgAlaGlyArgLeuArgArgProArgArg 520
QY 1114 TGC CGCGCTTCATTCGCGTACGTCAAATACCGAAACCGCTGCTCAAAGCGGCATCAA 1173
Db 521 GlyGlyArgGlyProGlyLeuGlnGlyAspProGlyVal----- 533
QY 1174 ACTTACGAGTGCACCAACCATGCGCTCCCGCCCAAAAGCAAGAGCGCTGACCGG 1233
Db 534 -----ArgProGluArgAspArgArgProGlyGln 544
QY 1234 CAGTCCGTACACGCGCTGCATGCCAAACCTTCATTGTGGACGCGCAACGCTTCAT 1293
```

```
Db 545 GlyProArgThrAspAlaThrAlaProGlyAlaGlyGlyArg----- 559
QY 1294 CGGCTCATTTAACCTCGACCCCGTTCGCGACGGTCAATACCGAAATGGCGTGTGTCAT 1353
Db 560 -----GlyAspProGlyAlaArgHisGlnHisLeuHisGlyTyArgAsp 574
QY 1354 CGAAAGCCCCAAATCGCAGAACAGATGGAGCGGACCCCTCGCGA-----TACCAC 1404
Db 575 GlnArgHisGlnArgValArgGlyThrGlyArgGluArgArgArgAlaAspArgHis 594
QY 1405 ACCCGAATACGCCTA-----CCGGTTACCTCGACAAACCAACCGCT 1449
Db 595 LeuAlaAlaArgLeuIleAlaGlyLeuSerArgArgGlyArgCysValArgHisPro 614
QY 1450 GCAATGGCACGATCCCGCCACCGCAAAACCTTACCCGAAACGA 1491
Db 615 AlaProGly---GluArgProProGluArgLeuThrGluArg 627

RESULT 37
US-09-252-991A-30396
; Sequence 30396, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 30396
; LENGTH: 694
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (334)
; OTHER INFORMATION: Identity of amino acid at the above locations are unknown.
US-09-252-991A-30396

Alignment Scores:
Pred. No.: 1,32e-12 Length: 694
Score: 263.00 Matches: 168
Percent Similarity: 34.5% Conservative: 43
Best Local Similarity: 27.5% Mismatches: 216
Query Match: 9.2% Indels: 185
DB: 2 Gaps: 34

US-10-665-990A-13 (1-1561) x US-09-252-991A-30396 (1-694)
QY 111 AAGACGACGAGCGAAGCGCTCATTTCAATCTTCCAAACCTGCTCTCTGACACATCC 170
Db 17 ArgArgGlyProAlaAlaArgIleAlaMetValPro-AlaAlaThrProGlyGln----- 34
QY 171 TCGAAATCCGCGACACCCCTCATATAACACGGGCT---ATCCGACATCTACCTGCTCGAGC 227
Db 35 -----GlyHisGlyHisArgArgArgGlyGlnLeuArgHis---ArgAlaArgArg 50
QY 228 ACCCGCCAGAGCCCTTCCCGCGCGCGCCCTTATCGAATCTGCCAAACACAGCCCTCG 287
Db 50 G-----ProLeuArgPro---ArgProGlyArgArgValArgLeuGlnArg 65
QY 288 ATTTGCAATACTATTGCGCAACGACATTTCCGCGAGGCTGCTGTTCACCTCATGT 347
Db 65 gValArgLeu-----ArgProAlaAlaArgProAspPr 77
QY 348 ACCTTGGCGCAGAACGCGCGCTGCGGTACGCCT---GCTGTGGACGACCAACACAGC 404
```

```
Db 77 oProGlyValArgHisGlyArgAlaGlnArgProAlaAlaAlaGlyAlaGlnGlyHisG1 97
Qy 405 GCGGTGGACGATCTCTGCTCGCTCGCTCGACAG----- 438
Db 97 yArgLeuProGluSerProArgArgProArgGlnLeuValValHisValLeuLeuGlnAr 117
Qy 439 -----CCATCCCAATAT-----CGAAGTGGCGCTGTTCACCC 470
Db 117 gHisLeuArgArgLeuProArgProGlyGlnHisProAlaArgLeuPhePro--ArgPr 136
Qy 471 CTTTCGCTCCTACGAAATGGCGGCACCTCGCTACCTGCGAGTTCCTCCCGCTCAACC 530
Db 136 oValArgProGluProGlyArgArgLeuLeuHisArgArgLeuArgLeuArgG1 156
Qy 531 GCCCATGTCACAAATCTTTACCGCGCAGAACCGCGCAC---CATACTGGCGGAC 587
Db 156 nProAspAlaProla-----ArgTrpArgProArgSerHisArgArgHisPr 173
Qy 588 GCAATATCGGCGACGAATACTTCAAGTCGTGAGGACACCGT----- 630
Db 173 oGlnProAlaAspGlyValHisProGlyArgHisLeuHisArgArgGlyArgPheProLe 193
Qy 631 -TTTCGGCGCCTGACATCTCCACCGCGCAGCTGTCGGCGAAGTATCGCACGACT 689
Db 193 uValGlnArgAspGlyArgProArg---ProValArgGlyArgHisAlaGlnSerArgCy 212
Qy 690 TCGACCGCTACTGGGCAAGCCATTC----- 714
Db 212 sArgGlnTrpArgGlyValProValGlyThrAlaAlaLeuProGlnGlyAsnArgArgAs 232
Qy 715 ----CGCCCAACAGCCACCGCATCATCCGACGCGCAACATCGGCAAGGCTTCAAG 770
Db 232 pAspArgAlaAspArgHisGly-----ArgArgHisArgArgLeuLeuProAs 248
Qy 771 CACTCGGATCAACGA-----CGAATCATCCAGACACGCGCTCTCGGCT 815
Db 248 pCysArgProArgArgAspGlnAlaLysHisArgArgLeuProAlaArgProValAlaVa 268
Qy 816 ACCGCGAAACGTCGACAGTCGCGCCCTCTACCAAAAAAATACAGACGG-----ACGCA 869
Db 268 lArgGlnSerArgArgAlaGlyLeuValArgProProArgGlyGlnAlaAlaLeuAlaHi 288
Qy 870 TCGACTGGCAGAGCGTCAAAACCGCTGATCAGCGACACCCCTGCAAAAGGACTCGACC 929
Db 288 sHisLeuGlyAlaArgProArgSerPro-----ProArgGlySerG1 302
Qy 930 GCGACCGCCGCAACCGCCGATTCGCGGAGGTGCAAGACGCGCTCAAAACAGCCGAAA 989
Db 302 uArgGlyAlaAlaThrGluHisTrpArgGlyGlnArgLeuArgAlaSerHisGlyLysPr 322
Qy 990 AAAGCGTCTATCGTTTCCCTTATTTTCGTCCTTACAAATCCGGCACAGACGACTGG 1049
Db 322 oGlyCysAlaAlaGly-----GlyHisProGlyAlaGly***GlyArgGlnGlnG1 339
Qy 1050 CAAAACTGGTCGAGGCGCATGCTTACCGTCT----- 1086
Db 339 y---ThrProAlaGlyHisArgArgArgCysGlnProValArgArgTrpArgProGlyG1 358
Qy 1086 ----- 1086
Db 358 yAlaGlnGlnProAlaSerAlaGlyProGlyLeuLeuCysAspSerGlyAsnLeuGlyHi 378
Qy 1087 -----GACCAACTCGCTACAGCGCGACGCG 1112
Db 378 sArgAlaValProGlyProProAlaAlaGlyAlaGluProLeuAlaThrGlyGlnArg-- 397
Qy 1113 TTGCGCGCTCCATTCGCGGTAGCTCAATACCGAAAAACCGCTGCTCAAAAGCGCGATCA 1172
Db 398 ---ArgArgPro-----AlaAlaAlaHisHisProHisAr 408
Qy 1173 AACTCTACGAGTGCACCCCAACATGCGCTCCCGCCACAAAAGCAAAAGGCGCTGAC-- 1230
Db 408 gPro---GlyAlaAlaArgHisProLeuHisProGlyProArgArgArgLeuProGlyLe 427
```

```
Qy 1231 -----CGCGAGCTCGTAACCGCTGCATGCCRAAACCTTCATTGTGCGCGCAAC 1283
Db 427 uProValAlaArgArgValAlaProAlaAspGlnArgProArgLeuGlyThrAlaAl 447
Qy 1284 GCATCTTCATCGGCTCATTCACACTCGACCC----- 1314
Db 447 aHisAlaThrArg---AlaAspProArgProGlyThrGlyProAlaSerGlyGlyArgLe 466
Qy 1315 -----CCGTTCCGACCGCTCAATACCGA----- 1338
Db 466 uSerGlyArgArgSerGlyAlaArgGlnProValAlaThrGlyGlnArgArgLeuG1 486
Qy 1339 -AATGGCGCTCATCGAAGCCCAAAAT-----CGCAGACAGATCGAGCGCACCC 1391
Db 486 yAspAlaGlyArgGlnArgHisProGlnHisProAlaGlnArgGlyAlaAlaG1 506
Qy 1392 TCCCGCATACACACCGCATACGCTTACCG---CGTTACCTCCGACAAACACCAACCGCC 1448
Db 506 nArgArgGlyProGlySerArgArgProProGlyArgGlnProGlyGlnArgGlnArgPr 526
Qy 1449 TGCATGGCGACGATCC-----CGCCACCGCAAAACCTACCCCGAAGCAACCC-- 1494
Db 526 oAlaGlyThrArgGlySerArgAlaArgArgGlnProArgArgArgProArgAlaThrG1 546
Qy 1495 -----CGAAGCGCAAACTTTGGAACGCGAT----- 1518
Db 546 yProLeuAlaThrAlaAlaThrAlaGluAlaArgProAlaLeuArgArgLeuAlaGlyPr 566
Qy 1519 -----CGCCGCAAAATCTATCCCTGCT 1542
Db 566 oAlaValAlaArgArgIleProProValProAla 577
```

RESULT 38

```
US-09-252-991A-30397
; Sequence 30397, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 30397
; LENGTH: 894
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-30397
```

```
Alignment Scores:
Pred. No.: 1,59e-12 Length: 894
Score: 262.50 Matches: 173
Percent Similarity: 32.3% Conservative: 47
Best Local Similarity: 25.4% Mismatches: 199
Query Match: 9.2% Indels: 263
DB: 2 Gaps: 40
```

```
US-10-665-990A-13 (1-1561) x US-09-252-991A-30397 (1-894)
```

```
Qy 160 GGACAAATCTCTGCAATACCGGCACACCCCTCATATAACCGGCT----- 204
Db 220 GlyGlnArgProGlnAlaThrArgHisGlyArgAspArgArgAlaProLeuArgLeuAla 239
Qy 205 -----ATCCGACATCTACCTGCTCGAGA-----CCCCCA 234
Db 240 AlaGlyThrProThrGlyProHisArgArgGlnThrProAlaArgArgAlaGlySerPro 259
```


RESULT 40

US-09-252-991A-25361
; Sequence 25361, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 25361
; LENGTH: 1037
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-25361

Alignment Scores:

Pred. No.:	2,02e-12	Length:	1037
Score:	261.50	Matches:	182
Percent Similarity:	33.0%	Conservative:	42
Best Local Similarity:	26.8%	Mismatches:	191
Query Match:	9.2%	Indels:	263
DB:	2	Gaps:	42

US-10-665-990A-13 (1-1561) x US-09-252-991A-25361 (1-1037)

QY	145	CAAACTGCTCTCGGCAACATCTCGAATTCGGCACACCC-----	189
DB	144	GlnAlaAlaArgProGlyGln---ProAlaGluAspArgProValAlaGlyGluPro	162
QY	190	---TCATAACACGGCTATCCGACAT-----CTACCTGCTCGACGACCCCA	234
DB	163	GlyAspArgGlnArgGlnSerHisGlyAlaAlaGluLeuProTrpArgAspProGlu	182
QY	235	CGAAGCCCTTGC---CGCCCGCGCCCTTATCGAATCTCGCCGAACA-----	279
DB	183	GluLeuProPheProArgAlaArgThrGlyArgArgLeuArgArgSerHisProProAla	202
QY	280	-----CAGCTCGATTGCAATACATATTGGCGCAACGACATTTCCGGCAGGCT	330
DB	203	GlnProGluArgProGlnSerHisGlnArgLeuArgGlnArgHis-----	218
QY	331	GCTGTTCAACCTCATCTCTCCGCGAGAACGCGCGTGGCGTACGCTGCTGTGGA	390
DB	219	-----GlnGlyArgTyrArgArgSerArgLysProAla-----	230
QY	391	CGACAAACACGCGGGTGGACGATCTCTGCTCGCCCTCGA-----	435
DB	231	-----HisArgLeuAlaGlyArgGlnLeuGlnArgProArgArgGlnProGlyAsp	247
QY	436	-----CAGCCATCCCAATATCGAATGCGCCTGTT	465
DB	248	AlaLeuGlnArgProGlyGlyTyrLeuGluProAlaLeuProArgProGlnProAla	267
QY	466	-----CAACCCCTTCTGCTACGAAATGGCGGCACCTCGGTACCTGACCGA	513
DB	268	ValAlaAlaPheGlnProIleAspProAspAlaLeuAla-----AlaLeuLeuGlnAspPro	286
QY	514	CTTCCCGCCCTCAACCG-----CGCATGCAACAAATCTTTTACCGCCGA-----	561
DB	287	ArgProGlyProGlnProAlaArgProAspArgGlnArgProLeuGlnArgArgAlaAsp	306
QY	562	-----CAACCG-----	567
DB	307	ProGlyArgArgProValAlaGluProGlnProAlaLysGlyArgTrpArgHisPro	326
QY	568	-----CGCCACCATACT	579

DB	327	ValArgLeuHisArgAlaGluArgProAspLeuProLeuArgLysGlnArgHisAlaVal	346
QY	580	CGCGGACGCAATATCGCGACGAATA-----	606
DB	347	ArgArgLeuGlnArgProArgAlaLeuValGluAspProValProGlyGlnProGly	366
QY	607	CTTAAAGTCGGTAGGACACCGTTTTCGCGCA-----	639
DB	367	LeuProArgProGlyLeuAlaValArgArgLeuArgProTrpProGlyArgPro	386
QY	640	---CCTGGACATCTCCCGCACCGCGCGCTCT-----	669
DB	387	AspProGlyArgProGlyGlnArgGlyLeuArgLeuProLeuGlnProGlnTrpGlnGlu	406
QY	670	-----CGCGAAGTATCGCACGACTTCCA-----	693
DB	407	ArgProAlaLeuGlyThrArgProValAlaAlaLeuArgLeuProArgArgProGlyGln	426
QY	694	-----CCGCTACTGGCAAGCATTCGCCCAACAGCCGCGCATCAT	738
DB	427	GlyProGlnArgAspProAlaLeuGlyHisPro---ProProArgArgArgLeuHisArg	445
QY	739	CGCAGCGCAACATCGCAAGGGTCTTCAAGCACTCGGATACCAACAGCAACATCCAG	798
DB	446	ProArgGlnHisProArgGln-----LeuGlnLeuArg-----	456
QY	799	ACACGCGCTCTCGCTACCGCAACCGTCGAACAGTCGCCCTCTTACCAAAAATACA	858
DB	457	---ArgIleProGlyGlyArgLeuProAspGlnValAlaLeuAsnProArgAsn---	474
QY	859	GACGGAGCATCGACTGCGAGCGCTCCAAACCGCCTGATCAG---CGACACCCCTGC	915
DB	475	GlnLysLysGluLysLeuProHisAlaCysSerProAlaThrHisLeuArgArgProLeu	494
QY	916	AAAGGACTCGACCGCGCAACCGCGGATTCGCGGAGGCTGCAAGACGCGCT	975
DB	495	ArgArgThrLeuProGlyGlyProArg-----ArgArgArgAlaAlaAla	510
QY	976	CAACAGCCGAAAAAGGCTATCTGTGTTTACCCCTATTTCCTTACAAAATCCCG	1035
DB	511	AlaGlyAlaArgGlnArgGlnHisGly-----ArgArgGlnProSerArg	526
QY	1036	CACAGA---CGCAGCTGCG-----AAACTGCTGCGAGCGCATAGAGTACCGT	1083
DB	527	HisAlaLeuArgThrAlaProAlaIleArgSerGlyLysGlnArgGlnArgArgArg	546
QY	1084	CCT---GACCAACTCGCTACAGC-----	1104
DB	547	AlaPheGlySerArgSerAlaThrGlySerTrpHisAlaValAlaAlaLeuArgLeu	566
QY	1105	-----GACCGAGTTGCGCGCTCCATTCGCGGTACGTCAATACCGAAAACCGCT	1155
DB	567	ProGlnProGluHisProAlaArgProLeuAlaArgArg-----SerAla	582
QY	1156	GCTAAAGCCGCGCATCAAACTCTACAGCTGCAACCCCAACCATCGCTCCCGCCACAAA	1215
DB	583	AlaAlaAlaGlyThrProGlyProArgTyrArgArgProGlyArgProAlaHisPro	602
QY	1216	AGACAAAGCCCTGACCGCGAG-----	1236
DB	603	ProGlyArgProAspArgGlnSerAlaAspAlaAlaMetGlyAlaArgTyrAlaValArg	622
QY	1237	-----CTCCGTAAACCGCTGCGATGCCAAAACCTTCATTGTT	1272
DB	623	ProProLeuAlaAlaProGluLeuArgLeuLeuProAlaHisArgArgLeuArgGln	642
QY	1273	GGA-----CGCAACCGCATCTTCATCGCTCATTCACCT-----	1308
DB	643	GlyArgArgAspGlnHisArgGlnGlnHisLeuAsp-----LeuGlnProAlaLeuArg	660
QY	1309	-----CGACCCCGCTTCCGACGCGCTCAA	1332

```
Db      661 LeuHisLeuGlyThrGlyArgAlaProGlyThrGlnArgProPro-----Ala 676
QY      1333 TACCGAAATGGCGTCGTCATCGAAAGCCC-----CAAATCGCAGACAA 1377
Db      677 LeuArgLeuAspGlnProGlnArgProGlnArgProLeuAlaProGlyArgLysArgArgThr 696
QY      1378 GATGGA-----GCGCACCCCTCGCCGATAC 1401
Db      697 AlaGlyThrGlyAlaAlaArgGluLeuLeuAspLeuLeuArgAlaGlyProAlaLeuAla 716
QY      1402 CACACCCGAATACGCCTACCG-----CGTTACCTCGACAAACACAAACCG 1446
Db      717 AspArgProGlyArgLeuProValGluThrAspGlnArgProHisArg---ArgPro 735
QY      1447 CCTGCA-----ATGGCAGCATCC-----CGCCACCCCGAAAAAC-----CTACCC 1485
Db      736 ProAlaGlyArgLeuAlaArgAlaGlyLeuArgHisArgProGlyGlyAspValGlnPro 755
QY      1486 GAACGAACCCGAAGC-----CAAACCTTTGGAACGCGATCGCCGCAAAATCCT 1533
Db      756 GlyGlnAlaAsnAlaValArgGlnProLeuArgLysArgArgAlaGlnPro 773
```

Search completed: May 2, 2006, 05:28:41
Job time : 146.5 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: May 2, 2006, 05:01:11 ; Search time 51.6 Seconds
(without alignments)
2658.412 Million cell updates/sec

Title: US-10-665-990A-13

Perfect score: 2852

Sequence: 1 caaatacaggaatgcgcgt.....tgcccatcgagggtttatta 1561

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 4886326

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+n2p.model -DEV=xlp
-Q=/abss/ABSSWEB spool/US10665990/runat_01052006_111938_9478/app_query.fasta_1
-DB=A_Geneseq -QFMT=fastan -SUFFIX=rag -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=40 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HFAPEXT=500 -MINLEN=0 -MAXLEN=200000000 -HOST=abss02p
-USER=US10665990 @CGN 1 1 348 @runat_01052006_111938_9478 -NCPU=6 -ICPU=3
-NO_MMAP -NRG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

A_Geneseq 21:.*
1: geneseq1980s:.*
2: geneseq1990s:.*
3: geneseq2000s:.*
4: geneseq2001s:.*
5: geneseq2002s:.*
6: geneseq2003as:.*
7: geneseq2003bs:.*
8: geneseq2004s:.*
9: geneseq2005s:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2681	94.0	525	ABP79466	Abp79466 N. gonorr
2	2630	92.2	525	ABB78070	Abb78070 Amino aci
3	2606.5	91.4	507	AA75751	Aay75751 Neisseria
4	2586	90.7	508	AA75753	Aay75753 Neisseria
5	2573	90.2	508	AA75752	Aay75752 Neisseria
6	831.5	29.2	543	ADL04719	Adl04719 M. catarr
7	790	27.7	564	ABO75163	AbO75163 Pseudomon
8	632	22.2	502	AAW55452	Aaw55452 H. pylori
9	632	22.2	502	AAW55550	Aaw55550 H. pylori

10	590	20.7	428	2	AAW98378	Aaw98378 H. pylori
11	493.5	17.3	271	6	ADA34272	Ada34272 Acinetoba
12	478.5	16.8	275	6	ADA34304	Ada34304 Acinetoba
13	421	14.8	2519	7	ABM88218	Abm88218 Rice abio
14	409	14.3	321	6	ADA36398	Ada36398 Acinetoba
15	336.5	11.8	5178	9	AEA04533	Aea04533 Human pro
16	336.5	11.8	5179	4	AAW24516	Aaw24516 C899p pre
17	336.5	11.8	5179	6	ABP55365	Abp55365 Human col
18	336.5	11.8	5179	6	ABO7258	AbO7258 Human p53
19	336.5	11.8	5179	7	ADD48091	Add48091 Human pro
20	336.5	11.8	5179	7	ADD44998	Add44998 Human pro
21	336.5	11.8	5179	8	ADQ29695	Adq29695 Human col
22	336.5	11.8	5179	8	ADQ80379	Adq80379 Intestina
23	336.5	11.8	5179	9	AEA27639	Aea27639 Human int
24	328	11.5	395	4	ABU53160	Abu53160 Human tes
25	326.5	11.4	822	7	ABO73174	AbO73174 Pseudomon
26	325	11.4	386	4	ABU53159	Abu53159 Human tes
27	325	11.4	717	4	ABU53144	Abu53144 Human tes
28	318	11.2	385	4	ABU53156	Abu53156 Human tes
29	314.5	11.0	387	4	ABU53161	Abu53161 Human tes
30	310.5	10.9	368	4	ABU53157	Abu53157 Human tes
31	304	10.7	513	6	ABM71816	Abm71816 Staphyloc
32	303.5	10.6	956	6	ABO53053	AbO53053 Human put
33	303.5	10.6	2296	8	ADN60280	Adn60280 Human spl
34	303.5	10.6	2752	8	ADN60278	Adn60278 Human ser
35	303.5	10.6	2752	8	ADP24102	Adp24102 PRO polyd
36	303.5	10.6	2752	9	ADY91619	Ady91619 Human pro
37	303	10.6	655	7	ABO68366	AbO68366 Pseudomon
38	302	10.6	618	7	ABO74950	AbO74950 Pseudomon
39	301.5	10.6	2263	7	ADJ70425	Adj70425 Human hea
40	301	10.6	618	7	ABO79612	AbO79612 Pseudomon
41	300.5	10.5	695	4	ABU53152	Abu53152 Human tes
42	300.5	10.5	745	4	ABU53154	Abu53154 Human tes
43	296	10.4	486	7	ADG30702	Adg30702 Xanthomon
44	295	10.3	507	7	ABO74852	AbO74852 Pseudomon
45	294.5	10.3	504	5	ABB49813	Abb49813 Listeria

ALIGNMENTS

RESULT 1

ABP79466

ID ABP79466 standard; protein; 525 AA.

XX

AC ABP79466;

XX

DT 07-MAR-2003 (first entry)

XX

DE N. gonorrhoeae amino acid sequence SEQ ID 5462.

XX

KW Antibacterial; infection; vaccine; gene therapy.

XX

OS Neisseria gonorrhoeae.

XX

FN WO200279243-A2.

XX

PD 10-OCT-2002.

XX

PF 12-FEB-2002; 2002WO-IB002069.

XX

PR 12-FEB-2001; 2001GB-00003424.

XX

PA (CHIR-) CHIRON SPA.

XX

FI Fontana MR, Pizza M, Masignani V, Monaci E;

XX

DR WPI; 2003-058415/05.

XX

DR N-PSDB; ABZ40436.

XX

PT New protein from Neisseria gonorrhoeae, useful for the manufacture of a

XX

PS medicament for treating or preventing N. gonorrhoeae infection.

PS

PS

PS

PS

XX

The present invention relates to proteins from *Neisseria gonorrhoeae*.
 CC Also disclosed are the nucleic acid molecules encoding the proteins and
 CC antibodies that specifically bind to the proteins. The composition
 CC comprising the protein, nucleic acid or antibody is useful for the
 CC manufacture of a medicament for treating or preventing *N. gonorrhoeae*
 CC infection, this may be in the form of a vaccine or gene therapy.
 CC Sequences given in records ABP76736-ABP81046 represent nucleic acid
 CC molecules of the invention
 XX

SQ Sequence 525 AA;

Alignment Scores:

Pred. No.: 9,466-189 Length: 525
 Score: 2681.00 Matches: 518
 Percent Similarity: 99.8% Conservative: 1
 Best Local Similarity: 99.6% Mismatches: 1
 Query Match: 94.0% Indels: 0
 DB: 6 Gaps: 0

US-10-665-990A-13 (1-1561) x ABP79466 (1-525)

QY 2 AAAATACAGCAATCGCTGTAACATATATCCCGATGAAACACGACGCTCATTTCC 61
 DB 6 LysIleGluAlaMetProSerGluThrIleSerProMetLysThrArgSerLeuLeuSer 25
 QY 62 CTTTATGCTCCTCTCTGTTTCATGTTTCTCATGTTGCCCCCACTGGAAGACGGACG 121
 DB 26 LeuLeuCysLeuLeuLeuCysSerCysSerSerTyrProLeuGluGluArgThr 45
 QY 122 GAAAGCGTCATTTCATTAATCTCCAAACCTGTCCTCTGGACACATCTCGAATCCGG 181
 DB 46 GluSerArgHisPheAsnThrSerLysProValLeuLeuAspAsnIleLeuGlnIleArg 65
 QY 182 CACACCCCTCATAAACAGGGCTATCCGACATCTACTGCTCGACACCCCGACGAAGCC 241
 DB 66 HisThrProHisAsnAsnGlyLeuSerAspIleTyrLeuLeuAspAspProHisGluAla 85
 QY 242 CTTGCGCGCGCGCCCTTATCGAATCTGCGGAACACAGCGCTCGATTGCAATACTAC 301
 DB 86 PheAlaAlaArgAlaAlaLeuIleGluSerAlaGluHisSerLeuAspLeuGlnTyrTyr 105
 QY 302 ATTTGGCGAACGACATTTCCGCGAGCTGCTGTTCACCTCATGTACCTTCGCGAGAA 361
 DB 106 IleTyrArgAsnAspIleSerGlyArgLeuLeuPheAsnLeuMetTyrLeuAlaGlu 125
 QY 362 CGCGCGCTGCGGTACGCTGCTGTTGGACGACAAACACACGCGCGGTGGACGATCTC 421
 DB 126 ArgGlyValArgValArgLeuLeuLeuAspAspAsnThrArgGlyLeuAspLeu 145
 QY 422 CTGCTCGCCCTCGACAGCAATCCCAATATCGAAGTGGCTGTTCACCCCTTGTCTTA 481
 DB 146 LeuLeuAlaLeuAspSerHisProAsnIleGluValArgLeuPheAsnProPheValLeu 165
 QY 482 CGCAATGCGCGCACTCGCTACCTGACCGACTCCCGCTCAACCGCGCATGCGAC 541
 DB 166 ArgLysTrpArgAlaLeuGlyTyrLeuThrAspPheProArgLeuAsnArgMetHis 185
 QY 542 AACAAATCTTTACCGCGCACACCGCGCCACCATACTCGCGGAGCGCAATATCGCGGAC 601
 DB 186 AsnLysSerPheThrAlaAspAsnArgAlaThrIleLeuGlyGlyArgAsnIleGlyAsp 205
 QY 602 GAATATCTCAATCGGTGAGGACACCGTTTTCGCGACTGACATCTCGCCACCGCGC 661
 DB 206 GluTyrPheLysValGlyGluAspThrValPheAlaAspLeuAspIleLeuAlaThrGly 225
 QY 662 AGCGTCTGGCGAAGTATCGACGACTTCGACCGCTACTGGCAAGCCATTCGCGCCAC 721
 DB 226 SerValValGlyGluValSerHisAspPheAspArgTyrTrpAlaSerHisAlaHis 245
 QY 722 AAGCCACGCGCATCTATCCGCGACGCGCAACATCGGAAGGGTCTTCAAGCACTCGGATAC 781
 DB 246 AsnAlaThrArgIleIleArgSerGlyAsnIleGlyLysGlyLeuGlnAlaLeuGlyTyr 265

QY 782 AACGACGAAACATTCAGACACGCGCTCTCGCTACCGGAAACCGTCGAAACAGTCGCC 841
 DB 266 AsnAspGluThrSerArgHisAlaLeuLeuArgTyrArgGluThrValGluGlnSerPro 285
 QY 842 CTCTACCAAAAATACAGACGCGGCGCATCGACTCGAGAGCGTCCAAACCGCTGATC 901
 DB 286 LeuTyrGlnLysIleGlnThrGlyArgIleAspTrpGlnSerValGlnThrArgLeuIle 305
 QY 902 AGCGACACCCCTGCAAAAAGGACTCGACCGCGACCGCCGCAACACCGCCGATTCGCGGAGG 961
 DB 306 SerAspSerProAlaLysGlyLeuAspArgAspArgLysProProIleAlaGlyArg 325
 QY 962 CTGCAAGACGCGCTCAAAACAGCCCGAAAAAGCGTCTATCTGGTTTACCCCTATTTCGTC 1021
 DB 326 LeuGlnAspAlaLeuLysGlnProGluLysSerValTyrLeuValSerProTyrPheVal 345
 QY 1022 CCTACAAAATCGGCGACAGACGCTGCGCAAACTGGTGCAGACGGCATAGACGTTACC 1081
 DB 346 ProThrLysSerGlyThrAspAlaLeuAlaLysLeuValGlnAspGlyIleAspValThr 365
 QY 1082 GTCTTGACCAACTCGCTACAGGCGACCGTTCGCGCGTCCATTTCGCGTACGTCAAA 1141
 DB 366 ValLeuThrAsnSerLeuGlnAlaThrAspValAlaAlaValHisSerGlyTyrValLys 385
 QY 1142 TACCGAAAACCGCTCTCAAGCCGCGCATCAAACTCTAGGACTGCAACCCGATGCC 1201
 DB 386 TyrArgLysProLeuLeuLysAlaGlyIleLysLeuTyrGluLeuGlnProAsnHisAla 405
 QY 1202 GTCCCGCGCACAAAAGAGAAAGGCGCTGACCGGCGCTCGTAACAGCGCTGCATGCCAAA 1261
 DB 406 ValProAlaThrLysAspLysGlyLeuThrGlySerSerValThrSerLeuHisAlaLys 425
 QY 1262 ACCTTTCATTGAGCGCAACCGCATCTTCATCGGCTCATCACTCGACCCCGCTTCC 1321
 DB 426 ThrPheIleValAspGlyLysArgIlePheIleGlySerPheAsnLeuAspProArgSer 445
 QY 1322 GCACGGCTCAATACCGAAATGGCGTCTCATCGAAAGCCCAAAATCGCAGAACAGATG 1381
 DB 446 AlaArgLeuAsnThrGluMetGlyValValIleGluSerProLysIleAlaGluGlnMet 465
 QY 1382 GAGCGCACCCCTCGCGATACACACCGCAATACGCTACCGGTTTACCTCGACAAACAC 1441
 DB 466 GluArgThrLeuAlaAspThrThrProGluTyrAlaTyrArgValThrLeuAspLysHis 485
 QY 1442 AACCCCTGCAATGCGACGATCCCGCCACCCGAAAAAATCTACCCGAAACGAAACCGGAGCC 1501
 DB 486 AsnArgLeuGlnTrpHisAspProAlaThrArgLysThrTyrProAsnGluProGluAla 505
 QY 1502 AAACCTTGGAAACGATCGCGCAAAAATCCCTATCCCTGCTGCCATCGAAGGTTTATTA 1561
 DB 506 LysLeuTrpLysArgIleAlaAlaLysIleLeuSerLeuLeuProIleGlyLeuLeu 525

RESULT 2
 ABB78070
 ID ABB78070 standard; protein; 525 AA.
 XX
 AC ABB78070;
 XX
 DT 29-AUG-2003 (revised)
 DT 05-NOV-2002 (first entry)
 XX
 DE Amino acid sequence of p55 polypeptide.
 XX
 KW p177; p88; p64; p55; p46; vaccine; gonorrhea.
 XX
 OS *Neisseria gonorrhoeae*.
 XX
 PN WO200260936-A2.
 XX
 PD 08-AUG-2002.
 XX
 PF 31-JAN-2002; 2002WO-US002881.

XX 31-JAN-2001; 2001US-0266070P.
 PR 06-AUG-2001; 2001US-0310356P.
 PR 23-OCT-2001; 2001US-0344452P.
 XX
 PA (IOWA) UNIV IOWA RES FOUND.
 PA (REGC) UNIV CALIFORNIA.
 PA (APIC) APICELLA M A.
 PA (EDWA) EDWARDS J L.
 PA (GIBS) GIBSON B W.
 PA (SCHE) SCHEFFLER K.
 PA (BROW) BROWN E.
 XX
 PI Apicella MA, Edwards JL, Gibson BW, Scheffler K, Brown E;
 XX WPI; 2002-619227/66.
 DR N-PSDB; ABQ78301.
 XX
 PT New polypeptide comprising p177, p88, p64, p55 or p46 from *Neisseria*
 PT gonorrhea, useful for preventing, or protecting a female patient
 PT against, *N. gonorrhea* colonization or infection.
 XX
 PS Claim 7; Page 120-121; 130pp; English.
 XX
 CC The present sequence represents a p55 polypeptide. The specification
 CC describes p177, p88, p64, p55 and p46 polypeptides from *Neisseria*
 CC gonorrhea. The polypeptides are useful as vaccines, for preventing, or
 CC protecting a female patient against, *N. gonorrhea* colonization or
 CC infection. Such immunization can prevent gonorrhea in women. (Updated on
 CC 29-AUG-2003 to standardise OS field)
 XX
 SQ Sequence 525 AA;

Alignment Scores:
 Pred. No.: 5, 4e-185 Length: 525
 Score: 2630.00 Matches: 509
 Percent Similarity: 98.7% Conservative: 4
 Best Local Similarity: 97.9% Mismatches: 7
 Query Match: 92.2% Indels: 0
 DB: 5 Gaps: 0

US-10-665-990A-13 (1-1561) x ABB78070 (1-525)

Qy 2 AAAATACAGCGAATCCGCTCTGAACATATATCCCGATGAAACACACGAGCTCATTTCC 61
 |||
 Db 6 LysThrGlnAlaMetProSerGluThrIleSerLeuMetLysThrArgSerLeuIleSer 25
 |||
 Qy 62 CTTTATGCTCTCTCTGTTCAATGTTCTTTCATGTTGCCCCCACTGGAGAACGAGC 121
 |||
 Db 26 LeuLeuCysLeuLeuLeuCysSerCysSerSerTrpLeuProLeuGluGluArgThr 45
 |||
 Qy 122 GAAAGCGTCAATTTCAATCTTCAACCTGCTCTCTGGACACATCTCTGCAATCGG 181
 |||
 Db 46 GluSerArgHisPheAsnThrSerLysProValArgLeuLeuAspAsnIleLeuGlnIleArg 65
 |||
 Qy 182 CACACCCCTCATAAACAGGCTATCCGACATCTACCTGCTCGACGACCCCGAGAGCC 241
 |||
 Db 66 HisThrProHisThrAsnGlyLeuSerAspIleTyrLeuLeuAsnAspProHisGluAla 85
 |||
 Qy 242 CTTGCCCGCGCGGCGCTTATCGAATCTGCGGAAACACAGCTCGATTGCAATATAC 301
 |||
 Db 86 PheAlaAlaArgAlaAlaLeuIleGluSerAlaGluHisSerLeuAspLeuGlnTyrTyr 105
 |||
 Qy 302 ATTTGGCGCAACGACATTTCCGGCAGGCTGCTGTTCACCTCATGTACCTTCCGCGGAA 361
 |||
 Db 106 IleTrpArgAsnAspIleSerGlyArgLeuLeuPheAsnLeuValTyrLeuAlaAlaGlu 125
 |||
 Qy 362 CGCGCGTGGCGTACGCTGCTTGGACGACAAACACGCGCGGTTGGAGATCTC 421
 |||
 Db 126 ArgGlyValArgValArgLeuLeuLeuAspAsnAsnThrArgGlyLeuAspAspLeu 145
 |||
 Qy 422 CTGCTCGCCTCGACAGCATTCCTATCGAAGTGGCTGTTCACCCCTTCGTCCTTA 481
 |||

Db 146 LeuLeuAlaLeuAspSerHisProAsnIleGluValArgLeuPheAsnProPheValLeu 165
 Qy 482 CGCAATGGCGCGCACTCGGCTACTGACCGACTTCCCGCCCTCAACCGCGCATGAC 541
 |||
 Db 166 ArgLysTrpArgAlaLeuGlyTyrLeuThrAspPheProArgLeuAsnArgArgMethis 185
 |||
 Qy 542 AACAAATCCTTTACCGCGCAACCGCGCCACCATCTCGCGGAGACGCAATATCGCGAC 601
 |||
 Db 186 AsnLysSerPheThrAlaAspAsnArgAlaThrIleLeuGlyGlyArgAsnIleGlyAsp 205
 |||
 Qy 602 GAATATCTTCAAGTCGTGAGACACCGCTTTTCCCGACCTCGGACATCTCTCCACCGCC 661
 |||
 Db 206 GluTyrPheLysValGlyGluAspThrValPheAlaAspLeuAspIleLeuAlaThrGly 225
 |||
 Qy 662 AGCGTCGCGCGAAGTATCGCAGACTTCGACCGCTACTCGGCAAGCCATTCGCGCCAC 721
 |||
 Db 226 SerValValGlyGluValSerHisAspPheAspArgTyrTrpAlaSerHisSerAlaHis 245
 |||
 Qy 722 AACGCCACCGCATCTCCGCGAGCGCAACATCGGCAAGGTCTTCAAGCACTCGGATAC 781
 |||
 Db 246 AsnAlaThrArgIleIleArgSerGlyAsnIleGlyLysGlyLeuGlnAlaLeuGlyTyr 265
 |||
 Qy 782 AACGACGAAACATCCAGACACCGCTCTCTGCGCTACCGCGAAACCGTCGAACTCGGCC 841
 |||
 Db 266 AsnAspGluThrSerArgHisAlaLeuLeuArgTyrArgGluThrValGluGlnSerPro 285
 |||
 Qy 842 CTCTACCAAAAATACAGAGCGGACGATCGACTCGGAGCGCTCCAAACCGCGCTGATC 901
 |||
 Db 286 LeuTyrGlnLysIleGlnThrGlyArgIleAspTrpGlnSerValGlnThrArgLeuIle 305
 |||
 Qy 902 AGCGACACCCCTTCAAAAAGGACTCGACCGGACCGCGCAACCGCGATTGCCGGGAGG 961
 |||
 Db 306 SerAspAspProAlaLysGlyLeuAspArgAspArgLysProProlleAlaGlyArg 325
 |||
 Qy 962 CTGCAAGACCGCTCAAAACAGCCGCAAAAAGCGTCTATCTGTTTCACTTTCGTC 1021
 |||
 Db 326 LeuGlnAspAlaLeuLysGlnProGluLysSerValTyrLeuValSerProTyrPheVal 345
 |||
 Qy 1022 CCTACAAAATCCGCGACAGACGCTCGGCAAACTGCTGAGGAGCGATAGAGTTACC 1081
 |||
 Db 346 ProThrLysSerGlyThrAspAlaLeuAlaLysLeuValGlnAspGlyIleAspValThr 365
 |||
 Qy 1082 GTCTGACCAACTCGCTTACAGGCGACCGAGTTCGCCCGCTCCATTCGGGTACGTCAAA 1141
 |||
 Db 366 ValLeuThrAsnSerLeuGlnAlaThrAspValAlaAlaValHisSerGlyTyrValLys 385
 |||
 Qy 1142 TACGAAAACCGCTGCTCAAAAGCGGATCAAACTCTACAGCTGCAACCCCAACCATGCC 1201
 |||
 Db 386 TyrArgLysProLeuLeuLysAlaGlyIleLysLeuTyrGluLeuGlnProAsnHisAla 405
 |||
 Qy 1202 GTCCCGCCACAAAAGAGCGCTGACCGGAGCTCCGTAACCGCTGCGATGCCAAA 1261
 |||
 Db 406 ValProAlaThrLysAspLysGlyLeuThrGlySerSerValThrSerLeuHisAlaLys 425
 |||
 Qy 1262 ACCTTCATTTGACGCGCAAAACCGATCTTCATCGGCTCATTCACCTCGACCCCGCTTC 1321
 |||
 Db 426 ThrPheIleValAspGlyLysArgIlePheIleGlySerPheAsnLeuAspProArgSer 445
 |||
 Qy 1322 GCGCGCTCAATACCGAAATGGCGCTGCTCATCGAAAGCCCAAAAATCGCAGAACAGATG 1381
 |||
 Db 446 AlaArgLeuAsnThrGluMetGlyValValIleGluSerProLysIleAlaGluGlnMet 465
 |||
 Qy 1382 GAGCGACCTCCCGCATACACACCGAATACCGCTACCGCTTACCTTCGCAACACAC 1441
 |||
 Db 466 GluArgThrLeuAlaAspThrSerProGluTyrAlaTyrArgValThrLeuAspArgHis 485
 |||
 Qy 1442 AACCGCTCGCAATGGCAGCATCCCGCACCCGAAAAACCTTACCGAACGAAACCGAGGCC 1501
 |||
 Db 486 AsnArgLeuGlnTrpHisAspProAlaThrArgLysThrTyrProAsnGluProGluAla 505
 |||
 Qy 1502 AAACTTTGGAAACGCAATCGCCCGCAAAAATCTATCCCTGCTGCGCATTCGAAGGTTTATTA 1561
 |||
 Db 506 LysLeuTrpLysArgIleAlaAlaLysIleLeuSerLeuLeuProIleGluSerLeuLeu 525
 |||


```

XX AC AAW55452;
XX DT 24-JUN-1998 (first entry)
XX DE H. pylori ORF 02aell612_33203250_cl_51 secreted protein.
XX KW Cytoplasmic; vaccine; prevention; treatment; infection; envelope;
XX KW identification; binding compound; bacteria; life cycle; activator;
XX KW inhibitor; duodenal ulcer disease; chronic gastritis; diagnosis.
XX OS Helicobacter pylori.
XX PN WO9737044-Al.
XX PD 09-OCT-1997.
XX PF 27-MAR-1997; 97WO-US005223.
XX PR 29-MAR-1996; 96US-00625811.
XX PR 02-APR-1996; 96US-00758731.
XX PR 25-OCT-1996; 96US-00736905.
XX PR 28-OCT-1996; 96US-00738859.
XX PR 06-DEC-1996; 96US-00761318.
XX PA (ASTR ) ASTRA AB.
XX PI Smith D, Alm RA;
XX DR WPI; 1997-503122/46.
XX DR N-PSDB; AAV24861.
XX PT Helicobacter pylori nucleic acid sequences and encoded polypeptide(s) -
XX PT useful in vaccines to treat or prevent H. pylori infection and for
XX PT diagnosis of H. pylori infection.
XX PS Claim 14,94; Page 659-660; 1145pp; English.
XX CC This sequence is a H. pylori secreted protein. The protein may be used in
XX CC a vaccine to prevent or treat H. pylori infection or to identify H.
XX CC pylori polypeptide binding compounds, useful as potential H. pylori life
XX CC cycle activators or inhibitors. The DNA and probes derived from it may be
XX CC used for the identification of H. pylori in a sample and the diagnosis of
XX CC H. pylori infection. Nucleic acid sequences complementary to the DNA act
XX CC as antisense sequences and can be used to prevent the translation of H.
XX CC pylori mRNA. Antibodies against the protein can be used in immunoassays
XX CC to evaluate the abundance and distribution of H. pylori-specific
XX CC antigens. The genomic sequence of H. pylori (ATCC 55679) was determined
XX CC from overlapping contigs generated by mechanically shearing the bacterial
XX CC DNA. The sequences were analysed for ORF of at least 180 nucleotides, and
XX CC the predicted coding regions defined by computer evaluation. To identify
XX CC likely H. pylori antigens for vaccine development, the amino acid
XX CC sequences predicted from various ORF were analysed for significant
XX CC homology to other known or exported membrane proteins. Having identified
XX CC and determined the sequences of interest, particular regions can be
XX CC isolated from H. pylori by PCR amplification for recombinant polypeptide
XX CC production, e.g. in E. coli hosts
XX SQ Sequence 502 AA;

Alignment Scores:
Pred. No.: 7.88e-38 Length: 502
Score: 632.00 Matches: 161
Percent Similarity: 51.7% Conservative: 93
Best Local Similarity: 32.8% Mismatches: 181
Query Match: 22.2% Indels: 56
DB: 2 Gaps: 17

US-10-665-990A-13 (1-1561) x AAW55452 (1-502)
QY 176 ATCCGGCACACCCCTCATACAAACGGGCTATCCGACATCTAC----- 217
Db 30 IleSerTyrAspProTyrThrThrThrIleGlySerLeuTyrAlaLysAsnLeuLysGlu 49

```

```

QY 218 -----CTGCTCAGCAGACCCCGCAGAGCGCTTGTCCCGCCGCGC 253
Db 50 AsnProLysHisSerAlaAlaIleLeuLeuGluAspGlyPheAspAlaLeuHisArg 69
QY 254 GCCGCCCTTATCGAATCTGCCGAACACAGCTCGATTGTCATATACATATTGGCGCAAC 313
Db 70 ValGlyLeuIleArgMetSerGlnLysSerIleAspMetGlnThrTyrIleTyrLysAsn 89
QY 314 GACATTTCCGGCAGGCTCTTCAACCTCATGTACCTTGGCGCAGAACCGCGCTGCGC 373
Db 90 AspLeuSerSerGlnValIleAlaLysGluLeuLeuAsnAlaAlaAsnArgGlyValLys 109
QY 374 GTACGCTGCTGTGGACAGACAACACCGCGCGGTGGAC-----GATCTC 421
Db 110 ValArgIleLeuLeuAspAsn-----GlyLeuAspSerAspPheSerAspIle 126
QY 422 CTGCTGCGCCCTCGACAGCATCCCAATATCGAAGTGCCTGTTCAACCCCTTCGTCCTA 481
Db 127 MetLeu---LeuAsnPheHisLysAsnIleGluValLysIlePheAsnProTyrTyrIle 145
QY 482 CGCAAAATGGCGCAGCTCGGCTAC-----CTGACCGAGCTTCCCGCGCTCAACCGC 532
Db 146 ArgAsn---LysGlyLeuArgTyrPheGluMetLeuAlaAspTyrGluArgIleLysLys 164
QY 533 CGCATGCACAACAATCTTTACCGCCGACACCGCGCCACCATACTCGGCGGAGCGCAAT 592
Db 165 ArgMetHisAsnLysLeuPheIleValAspAsnPheAlaValIleIleGlyArgAsn 184
QY 593 ATCGCGCAGCAATACTTCAAAAGTCGGTGAGACACCGTTTTCGCGAGCTCGACATCTCTC 652
Db 185 IleGlyAspAsnTyrPheAspAsnAspLeuAspThrAsnPheLeuAspLeuAspAlaLeu 204
QY 653 GCCACCGCGAGCGTCTCGGCGAAGTATCGCAGCACTTCGACCGCTACTGGCAGCCCAT 712
Db 205 PhePheGlyGlyValAlaSerLysAlaLysGluSerPheGluAsnTyrTrpArgPheHis 224
QY 713 TCCGCCACACGCCCGCATCATCCGAGC-----GGC 748
Db 225 ArgSerIleProValSer---LeuLeuArgThrHisLysArgLeuLysAsnValLys 243
QY 749 AACATCGCAGAGGCTTCAAGCATCTCGATACACAGCAACATCCAGACACCGGCTC 808
Db 244 GluIleAlaLysLeuHisGluLysIleIleSerAlaGluAspAlaAsnGluPheGlu 263
QY 809 CTGCGCTACCGGAAACCGTCGACAGTCGCCCTCTACCAA---AAATACAGCGGA 865
Db 264 LysLysValAsnAspPheIleGluArgPheGlnLysTyrGlnTyrProIleTyrGly 283
QY 866 CGCATCGAGTCGACAGGCTCCAAACCGCTGATCAGGACACCCCTGCAAAAGACTC 925
Db 284 -----AsnAlaIlePheLeuAlaAspLeuProAlaLys---Ile 295
QY 926 GACCGCGACCGCGCAACCGCGGAGTTCGCGGAGGCTGCAAGACGCGCTCAAAACAGCC 985
Db 296 AspThrProLeuTyrSer---ProIleLysIleAlaPheGluLysAlaLeuLysAsnAla 314
QY 986 GAAAAAAGCGTCTATCTGGTTTCACCTTATTTGCTCCCTTACAAATCCGGAACCGCTCA 1045
Db 315 LysAspSerValPheIleAlaSerSerTyrPheIleProGlyLysLysIleMetLysIle 334
QY 1046 CTGGCAAAATGCTGCAGGACGGCATAGACGTTCGGTCTCAGCAACTCGCTACAGGCG 1105
Db 335 PheLysAsnGlnIleSerLysGlyIleGluLeuAsnIleLeuThrAsnSerLeuSerSer 354
QY 1106 ACCGAGCTTGGCGCGCTCCATTCCGGCTACGTCAATACCGAAACCGCTGTCTCAAAGCC 1165
Db 355 ThrAspAlaIleValValTyrGlyAlaTrpGluAUGTyrArgAsnLysLeuValArgMet 374
QY 1166 GGCATCAAACTCTAGAGCTGCAACCCCAACCATTCGCTCCCGCCGACAAAGCAAGG 1225
Db 375 GlyAlaAsnValTyrGluIleArgAsnAspPhePheAsnArgGlnIleLysGlyArg--- 393

```



```
Db      264 LysLysValAsnAspPheIleGluArgPheGlnLysTyrGlnTyrProIleTyrTyrGly 283
QY      866 CGCATCGACTGCGCAGAGCGTCAAAACCCCGCTGATCAGCAGCACCCCTGCAAAAGGACTC 925
Db      284 -----AsnAlaIlePheLeuAlaAspLeuProAlaLys---Ile 295
QY      926 GACCGCGACCGCGAAACCGCGATTGCCGGAGGCTGCAAGACGCGCTCAAAAGCGCC 985
Db      296 AspThrProLeuTyrSer---ProIleLysIleAlaPheGluLysAlaLeuLysAsnAla 314
QY      986 GAAAAAGCGTCTATCTGTTTCAACCTATTTCGCTCCCTACAAAATCCGGCACACAGCA 1045
Db      315 LysAspSerValPheIleAlaSerSerTyrPheIleProGlyLysLysIleMetLysIle 334
QY      1046 CTGGCAAAACTGGTCAGACGGCATAGACGTTACGCTCACCACAACTCGCTACAGGCG 1105
Db      335 PheLysAsnGlnIleSerLysGlyIleGluLeuAsnIleLeuThrAsnSerLeuSerSer 354
QY      1106 ACCGACGTTGCGCGCTCATTCGCGGTACGTCAAAATACCGAAAACCGCTGCTCAAGCC 1165
Db      355 ThrAspAlaIleValValTyrGlyAlaTrpGluArgTyrArgAsnLysLeuValArgMet 374
QY      1166 GCATCAAACTCTACGAGCTGCAACCCACCACTGCGTCCCGCCACAAAAGACAAGGC 1225
Db      375 GlyAlaAsnValTyrGluIleArgAsnAspPheAsnArgGlnIleLysGlyArg--- 393
QY      1226 CTGACCGCGCAGCTCGGTAACCACTGATGCGCAAAACCTTCATGTTGGACGCAACGC 1285
Db      394 -----PheSerThrLysHisSerLeuHisGlyLysThrIleValPheAspAlaLeu 411
QY      1286 ATCTTCATCGGCTCATTTCAACTCGACCCCGCTTCGCGACGGCTCAATACCGAAATGGGC 1345
Db      412 ThrLeuLeuGlySerPheAsnIleAspProArgSerAlaTyrIleAsnThrGluSerAla 431
QY      1346 GTCTGTCATCGAAGCCCAAAATCCGAAACAGATGGAGCGCACCTCGCCCATACCACA 1405
Db      432 ValIlePheAspAsnProSerPheAlaLysArgValArgLeuSerLeuLysAspHisAla 451
QY      1406 CCGCAATACGCTACCGCTTACCCTCGCAAAACACACACCGCTGCAATGGCAGCATCC 1465
Db      452 ---GlnGlnSerTrpHisLeuValLeuTyrArgHis---ArgValIleTrpGlu----- 467
QY      1466 GCCACCGGAAAACCTAC-----CCGAAGAACCCGGAAGCCCAACTTTGGAAA 1513
Db      468 AlaThrGluGluGlyIleLeuIleHisGluLysAsnSerProAspThrSerPhePheLeu 487
QY      1514 CGCATCGCGCGCAAAATCCTATCCCTGCTGCC 1546
Db      488 ArgLeuIleLysGluTrpSerLysValLeuPro 498

RESULT 10
ID AAW98378 standard; protein; 428 AA.
XX AC AAW98378;
XX AC AAW98378;
XX AC AAW98378;
DT 31-MAR-1999 (first entry)
DE H. pylori GHPO 1375 protein.
XX KW GHPO protein; Helicobacter infection; gastroduodenal disease; gastritis;
XX KW peptic ulcer disease.
XX OS Helicobacter pylori.
XX PN WO9843478-Al.
XX PD 08-OCT-1998.
XX PF 01-APR-1998; 98WO-US006371.
XX XX 01-APR-1997; 97US-00833457.
XX PR
```

```
PR      24-JUN-1997; 97US-00881227.
PR      29-JUL-1997; 97US-00902615.
XX (INNR ) MERIEUX ORAVAX PASTEUR MERIEUX SERUMS.
PA (HUMA-) HUMAN GENOME SCI INC.
XX Kleanthous H, Al-Garawi A, Miller C, Tomb J, Oomen RP;
XX WPI; 1998-542293/46.
DR N-PSDB; AAX14097.
XX New isolated Helicobacter polynucleotides - used to develop products for
PT the diagnosis, prevention and treatment of Helicobacter infections and
PT gastrointestinal diseases.
XX Claim 8; Page 581-583; 2054pp; English.
CC This sequence represents a Helicobacter pylori GHPO protein of the
CC invention. The polypeptides can be used for preventing or treating
CC Helicobacter infections, and gastroduodenal diseases associated with
CC these infections, including acute, chronic, and atrophic gastritis, and
CC peptic ulcer diseases, e.g. gastric and duodenal ulcers. They can also be
CC used for the production of antibodies. The products can also be used for
CC detection and diagnosis
XX SQ Sequence 428 AA;

Alignment Scores:
Pred. No.: 9,4e-35 Length: 428
Score: 590.00 MatChes: 145
Percent Similarity: 50.1% Conservative: 80
Best Local Similarity: 32.3% Mismatches: 146
Query Match: 20.7% Indels: 78
DB: 2 Gaps: 14

US-10-665-990A-13 (1-1561) x AAW98378 (1-428)
QY      272 GCCGAACACAGCCTCGATTTCGAATACATTTGGCGCAAGCACATTTCCGGCAGGCTG 331
Db      2 SerGlnLysSerIleAspMetGlnThrTyrIleTyrLysAsnAspLeuSerSerGlnVal 21
QY      332 CTGTTCAACCTCATGTACTCTTCCGCGAAGACGCGCGTGGCTGCTGCTGTGTGGAC 391
Db      22 IleAlaLysGluLeuLeuAsnAlaAlaAsnArgGlyValLysValArgIleLeuLeuAsp 41
QY      392 GACAAACAACACGCGCGGTGGAC-----GATCTCTGCTCGCCCTCGACAGC 439
Db      42 AspAsn-----GlyLeuAspSerAspPheSerAspIleMetLeu---LeuAsnPhe 57
QY      440 CATCCCAATATCGAAGTCGCGCTGTCAACCCCTTTCGTCCTACGCAAAATGGCGCACCTC 499
Db      58 HisLysAsnIleGluValLysIlePheAsnProTyrTyrIleArgAsn---LysGlyLeu 76
QY      500 GGCTAC-----CTGACCGACTTCCCGCCCTCAACCGCGCATGCACACAAATCC 550
Db      77 ArgTyrPheGluMetLeuAlaAspTyrGluArgIleLysLysArgMetHisAsnLysLeu 96
QY      551 TTTACCGCGCAACACCGCGCACCATACTCTCGCGGACGCAATATCGGCGCAACAAATCTTC 610
Db      97 PheIleValAspAsnPheAlaValIleIleGlyIleGlyArgAsnIleGlyAspAsnTyrPhe 116
QY      611 AAGTCGCTGAGGACACCGTTTTCGCCGACCTCGGACATCCTCGCACCGCGCAGCGTCGTC 670
Db      117 AspAsnAspLeuAspThrAsnPheLeuAspLeuAspAlaLeuPhePheGlyValAla 136
QY      671 GCGGAATGTCACGACTTCGACCGCTACTGGCAAGCCATTCGCCCCACCAACGCCACG 730
Db      137 SerLysAlaLysGluSerPheGluArgTyrTrpArgPheHisArgSerIleProValSer 156
QY      731 CGCATCTCCGCGCGGCAACATCGGCAAGGGTCTTCAAGCACTCGGATACACACGCA 790
Db      157 ---LeuLeuArgThr----- 160
```

```
QY 791 ACATCAGACACGCGCTCCTCGCTACCGCGAAACCGTCGACACGTCGCCCTCTACCAA 850
Db 161 -----HisLysArgLeuLys-----AsnAsnAlaLysGluAlaLysLeuHisGlu 176
QY 851 AAAATACAGACGGGACGATCGACTGGCAGACGTCCTCAAAACCGCGCTG----- 898
Db 177 LysIleProIleSerAlaGluAspLysAsnGlnPheGluLysLysValAsnAspPheIle 196
QY 899 -----ATCAGCGAC 907
Db 197 AspArgPheGlnLysTyGlnTyrProIleTyrTyrGlyAsnAlaIlePheLeuAlaAsp 216
QY 908 ACCCTGCAAAAGACTCGACCGCGACCGCCGCAAAACCGCGATTGCGGGAGGTGCAA 967
Db 217 SerProLysLys---IleAspThrProLeuTyrSer---ProIleLysIleAlaPheGlu 234
QY 968 GACGGCTCAACACGCCGAAAGAGGTCTATCTGGTTTACCCCTATTTCGTCCTTACA 1027
Db 235 LysAlaLeuLysAsnAlaLysAspSerValPheIleAlaSerSerTyrPheIleProgly 254
QY 1028 AAATCCGGCACAGACGCGACTGCAAAACTGGTGCAGGACGCGCATAGACGTTACCGTCGT 1087
Db 255 LysLysMetMetLysIlePheLysAsnGlnIleSerLysGlyIleGluLeuAsnIleLeu 274
QY 1088 ACCAACTCGTACAGCGACGACGCTTGC CGCGCTCCATTCCGGCTACGTCCTCAAAATACCGA 1147
Db 275 ThrAsnSerLeuSerSerThrAspAlaIleValTyrGlyAlaTyrGluArgTyrArg 294
QY 1148 AAACCGTGCTCAAGCGCGGATCAAACTCTACGAGCTGCAACCCCAACCATCGCGTCGCC 1207
Db 295 AsnGlnLeuValArgMetGlyAlaAsnValTyrGluIleArgAsnAspPheAsnArg 314
QY 1208 GCCACAAAGACAAAGGCTGACCGCGACGCTCGGTAACGACGCTGCATGCCCAAAACCTTC 1267
Db 315 GlnIleLysGlyArg-----PheSerThrLysHisSerLeuHisGlyLysThrIle 331
QY 1268 ATTGTGGACGCAACGACGCTTTCATCGGCTCATTCACCTCGACCCCGCTTCGCGACGG 1327
Db 332 ValPheAspAspAsnLeuThrLeuLeuGlySerPheAsnIleAspProArgSerAlaTyr 351
QY 1328 CTCATATACGAAATGGGGGTGTCTATCGAAAGCCCAAAATCGCGAAACAGATGAGCGC 1387
Db 352 IleAsnThrGluSerAlaValLeuPheAspAsnProSerPheAlaLysArgVal----- 369
QY 1388 ACCCTCGCGATACACACCGGATACGCTACGCGTACCGCTTACCCTCGACAAACACCGC 1447
Db 370 -----ArgLeuSerLeuLysAspHisAlaGln 378
QY 1448 CTGCAATGGCAGATCCCGCCACCGCAAAACCTACCGGAACGACCGCAACCAACTT 1507
Db 379 GlnSerTrpHisLeuValValTyrArg-----HisArgValIle 391
QY 1508 TGGAAACGCATCGCGGCAAAATCTTA 1534
Db 392 TrpGluAlaValGluGluGlyIleLeu 400
RESULT 11
ADA34272
XX AC
XX AD
XX AE
XX AF
XX AG
XX AH
XX AI
XX AJ
XX AK
XX AL
XX AM
XX AN
XX AO
XX AP
XX AQ
XX AR
XX AS
XX AT
XX AU
XX AV
XX AW
XX AX
XX AY
XX AZ
XX BA
XX BB
XX BC
XX BD
XX BE
XX BF
XX BG
XX BH
XX BI
XX BJ
XX BK
XX BL
XX BM
XX BN
XX BO
XX BP
XX BQ
XX BR
XX BS
XX BT
XX BU
XX BV
XX BW
XX BX
XX BY
XX BZ
XX CA
XX CB
XX CC
XX CD
XX CE
XX CF
XX CG
XX CH
XX CI
XX CJ
XX CK
XX CL
XX CM
XX CN
XX CO
XX CP
XX CQ
XX CR
XX CS
XX CT
XX CU
XX CV
XX CW
XX CX
XX CY
XX CZ
XX DA
XX DB
XX DC
XX DD
XX DE
XX DF
XX DG
XX DH
XX DI
XX DJ
XX DK
XX DL
XX DM
XX DN
XX DO
XX DP
XX DQ
XX DR
XX DS
XX DT
XX DU
XX DV
XX DW
XX DX
XX DY
XX DZ
XX EA
XX EB
XX EC
XX ED
XX EE
XX EF
XX EG
XX EH
XX EI
XX EJ
XX EK
XX EL
XX EM
XX EN
XX EO
XX EP
XX EQ
XX ER
XX ES
XX ET
XX EU
XX EV
XX EW
XX EX
XX EY
XX EZ
XX FA
XX FB
XX FC
XX FD
XX FE
XX FF
XX FG
XX FH
XX FI
XX FJ
XX FK
XX FL
XX FM
XX FN
XX FO
XX FP
XX FQ
XX FR
XX FS
XX FT
XX FU
XX FV
XX FW
XX FX
XX FY
XX FZ
XX GA
XX GB
XX GC
XX GD
XX GE
XX GF
XX GG
XX GH
XX GI
XX GJ
XX GK
XX GL
XX GM
XX GN
XX GO
XX GP
XX GQ
XX GR
XX GS
XX GT
XX GU
XX GV
XX GW
XX GX
XX GY
XX GZ
XX HA
XX HB
XX HC
XX HD
XX HE
XX HF
XX HG
XX HH
XX HI
XX HJ
XX HK
XX HL
XX HM
XX HN
XX HO
XX HP
XX HQ
XX HR
XX HS
XX HT
XX HU
XX HV
XX HW
XX HX
XX HY
XX HZ
XX IA
XX IB
XX IC
XX ID
XX IE
XX IF
XX IG
XX IH
XX II
XX IJ
XX IK
XX IL
XX IM
XX IN
XX IO
XX IP
XX IQ
XX IR
XX IS
XX IT
XX IU
XX IV
XX IW
XX IX
XX IY
XX IZ
XX JA
XX JB
XX JC
XX JD
XX JE
XX JF
XX JG
XX JH
XX JI
XX JJ
XX JK
XX JL
XX JM
XX JN
XX JO
XX JP
XX JQ
XX JR
XX JS
XX JT
XX JU
XX JV
XX JW
XX JX
XX JY
XX JZ
XX KA
XX KB
XX KC
XX KD
XX KE
XX KF
XX KG
XX KH
XX KI
XX KJ
XX KK
XX KL
XX KM
XX KN
XX KO
XX KP
XX KQ
XX KR
XX KS
XX KT
XX KU
XX KV
XX KW
XX KX
XX KY
XX KZ
XX LA
XX LB
XX LC
XX LD
XX LE
XX LF
XX LG
XX LH
XX LI
XX LJ
XX LK
XX LL
XX LM
XX LN
XX LO
XX LP
XX LQ
XX LR
XX LS
XX LT
XX LU
XX LV
XX LW
XX LX
XX LY
XX LZ
XX MA
XX MB
XX MC
XX MD
XX ME
XX MF
XX MG
XX MH
XX MI
XX MJ
XX MK
XX ML
XX MM
XX MN
XX MO
XX MP
XX MQ
XX MR
XX MS
XX MT
XX MU
XX MV
XX MW
XX MX
XX MY
XX MZ
XX NA
XX NB
XX NC
XX ND
XX NE
XX NF
XX NG
XX NH
XX NI
XX NJ
XX NK
XX NL
XX NM
XX NO
XX NP
XX NQ
XX NR
XX NS
XX NT
XX NU
XX NV
XX NW
XX NX
XX NY
XX NZ
XX OA
XX OB
XX OC
XX OD
XX OE
XX OF
XX OG
XX OH
XX OI
XX OJ
XX OK
XX OL
XX OM
XX ON
XX OO
XX OP
XX OQ
XX OR
XX OS
XX OT
XX OU
XX OV
XX OW
XX OX
XX OY
XX OZ
XX PA
XX PB
XX PC
XX PD
XX PE
XX PF
XX PG
XX PH
XX PI
XX PJ
XX PK
XX PL
XX PM
XX PN
XX PO
XX PP
XX PQ
XX PR
XX PS
XX PT
XX PU
XX PV
XX PW
XX PX
XX PY
XX PZ
XX QA
XX QB
XX QC
XX QD
XX QE
XX QF
XX QG
XX QH
XX QI
XX QJ
XX QK
XX QL
XX QM
XX QN
XX QO
XX QP
XX QQ
XX QR
XX QS
XX QT
XX QU
XX QV
XX QW
XX QX
XX QY
XX QZ
XX RA
XX RB
XX RC
XX RD
XX RE
XX RF
XX RG
XX RH
XX RI
XX RJ
XX RK
XX RL
XX RM
XX RN
XX RO
XX RP
XX RQ
XX RR
XX RS
XX RT
XX RU
XX RV
XX RW
XX RX
XX RY
XX RZ
XX SA
XX SB
XX SC
XX SD
XX SE
XX SF
XX SG
XX SH
XX SI
XX SJ
XX SK
XX SL
XX SM
XX SN
XX SO
XX SP
XX SQ
XX SR
XX SS
XX ST
XX SU
XX SV
XX SW
XX SX
XX SY
XX SZ
XX TA
XX TB
XX TC
XX TD
XX TE
XX TF
XX TG
XX TH
XX TI
XX TJ
XX TK
XX TL
XX TM
XX TN
XX TO
XX TP
XX TQ
XX TR
XX TS
XX TT
XX TU
XX TV
XX TW
XX TX
XX TY
XX TZ
XX UA
XX UB
XX UC
XX UD
XX UE
XX UF
XX UG
XX UH
XX UI
XX UJ
XX UK
XX UL
XX UM
XX UN
XX UO
XX UP
XX UQ
XX UR
XX US
XX UT
XX UV
XX UW
XX UX
XX UY
XX UZ
XX VA
XX VB
XX VC
XX VD
XX VE
XX VF
XX VG
XX VH
XX VI
XX VJ
XX VK
XX VL
XX VM
XX VN
XX VO
XX VP
XX VQ
XX VR
XX VS
XX VT
XX VU
XX VV
XX VW
XX VX
XX VY
XX VZ
XX WA
XX WB
XX WC
XX WD
XX WE
XX WF
XX WG
XX WH
XX WI
XX WJ
XX WK
XX WL
XX WM
XX WN
XX WO
XX WP
XX WQ
XX WR
XX WS
XX WT
XX WU
XX WV
XX WW
XX WX
XX WY
XX WZ
XX XA
XX XB
XX XC
XX XD
XX XE
XX XF
XX XG
XX XH
XX XI
XX XJ
XX XK
XX XL
XX XM
XX XN
XX XO
XX XP
XX XQ
XX XR
XX XS
XX XT
XX XU
XX XV
XX XW
XX XX
XX XY
XX XZ
XX YA
XX YB
XX YC
XX YD
XX YE
XX YF
XX YG
XX YH
XX YI
XX YJ
XX YK
XX YL
XX YM
XX YN
XX YO
XX YP
XX YQ
XX YR
XX YS
XX YT
XX YU
XX YV
XX YW
XX YX
XX YY
XX YZ
XX ZA
XX ZB
XX ZC
XX ZD
XX ZE
XX ZF
XX ZG
XX ZH
XX ZI
XX ZJ
XX ZK
XX ZL
XX ZM
XX ZN
XX ZO
XX ZP
XX ZQ
XX ZR
XX ZS
XX ZT
XX ZU
XX ZV
XX ZW
XX ZX
XX ZY
XX ZZ
```

13-MAY-2003.

04-JUN-1999; 99US-00328352.

09-JUN-1998; 98US-0088701P.

(GENO-) GENOME THERAPEUTICS CORP.

Breton G, Bush D;

WPI; 2003-576092/54.

N-PSDB; ADA30146.

New Acinetobacter baumannii proteins and nucleic acids, useful as reagents for diagnosing a bacterial disease, as components of antibacterial vaccines, as targets for antibacterial drugs, or as biocontrol agents for plants.

Example; SEQ ID NO 5559; 328bp; English.

The invention relates to isolated Acinetobacter baumannii nucleic acids. The A. baumannii nucleic acids and polypeptides are useful as reagents for diagnosing a bacterial disease, as components of antibacterial vaccines, as targets for antibacterial drugs, to detect the presence of A. baumannii and other Acinetobacter species in a sample, in screening compounds for the ability to interfere with the A. baumannii life cycle or to inhibit A. baumannii infection, and as biocontrol agents for plants. The present sequence represents the amino acid sequence of an A. baumannii protein.

Sequence 271 AA;

Alignment Scores:

Pred. No.: 1.08e-27 Length: 271

Score: 493.50 Matches: 102

Percent Similarity: 63.9% Conservative: 47

Best Local Similarity: 43.8% Mismatches: 73

Query Match: 17.3% Indels: 11

DB: 6 Gaps: 4

US-10-665-990A-13 (1-1561) x ADA34272 (1-271)

QY 59 TCCCTTTATGCTCCCTCTCTGTCATGTTCTTCTCATGTTGTCCTG----- 103

Db 37 SerSerPheAlaValAlaLeuThrGlyCysSerThr---LeuProLysHisSerProGlu 55

QY 104 CCATCTGGAAGACGACGCGAAGCCGTCATTTCATACTCTTCAACACCTGCTCTCTGAC 163

Db 56 ProIleGln-----TyrAlaArgAspIleAspThrSerGlnThrSerLeuSerLys 72

QY 164 AACATCTGCAATCCGCGACACCCCTCATACACGCGGCTATCCGACATCTACCTGCTC 223

Db 73 IleIleThrProLeuArgGlu-----LysAsnProAsnLeuThrGlyTyrHisLeuLeu 90

QY 224 GACGACCCCGACGAGCCCTTTCGCGCGCGCGCTTATTCGAATCTGCGAAGACACACG 283

Db 91 AsnAspProLeuGluAlaLeuAlaArgLeuArgLeuAlaGlyLysThr 110

QY 284 CTGATTGCAATACATATTGGCGCAACGACATTTCCGCGACGCTGCTGTTCACACCTC 343

Db 111 LeuAspLeuGlnTyrTyrIleTrpAspAsnAspLysValGlyAlaLeuAlaHisAla 130

QY 344 ATGTACCTTGGCGAGAACCGCGGCTGCGGTAGCTGCTGCTGTTGGAGCAGCAACACG 403

Db 131 IleIleArgAlaAlaAspArgGlyValLysValArgLeuLeuIleAspAspAsnAla 150

QY 404 CGCGGTTGGACGATCTCTGCTCGCCCTCGACGACCATCCCAATATCGAAGTCGCGCTG 463

Db 151 LysLysMetGluGlyValLeuLeuAlaLeuSerGlnHisLysAsnIleGluValLysLeu 170

QY 464 TTCACCCCTTCGTCCTACGCAATGCGCGCACTCGGCTACCTGACCGACCTCCCGCGC 523

Db 171 PheAsnProTyrArgPheArgLysTyrArgAlaMetAspMetIleLeuAspLeuLysArg 190

20-NOV-2003 (first entry)

Acinetobacter baumannii protein #1433.

Acinetobacter baumannii; bacterial disease; antibacterial; vaccine;

plant biocontrol agent.

Acinetobacter baumannii.

US6562958-B1.

QY 524 CTCACCGCGCATGCACAAATCTTTTACCGCGCAACACCGCGCCACCCATCTCGGC 583
Db ::
191 ILeAsnArgArgMetHisAsnLysSerPheIleAlaAspAsnGlnValAlaLeuIleGly 210
QY 584 GCACGCAATATCGGCGACGAATACTTCAAAGTCGGTGAGGACACCGTTTTCGCCACCTG 643
Db ::
211 GlyArgAsnMetThrAsnGlnTyrTyrAsnValSerAspSerTyrGlnPheSerAspVal 230
QY 644 GATATCTCCGCCACCGCAGCGTCGTCGGCGAAGTATCCACGCACTTCGACCGCTACTGG 703
Db ::
231 AspValMetLeuValGlyAlaAlaValAspAspIleValAsnSerPheValAspTyrTrp 250
QY 704 GCAAGCATTCGCCACACCAACGCCACGCGCATCATCCGC 742
Db ::
251 AsnHisGluTyrAlaTyrSerValGlnSerIleValArg 263

RESULT 12
ADA34304
ID ADA34304 standard; protein; 275 AA.
XX
AC ADA34304;
XX
DT 20-NOV-2003 (first entry)
XX
DE Acinetobacter baumannii protein #1465.
XX
KW Acinetobacter baumannii; bacterial disease; antibacterial; vaccine;
KW plant biocontrol agent.
XX
OS Acinetobacter baumannii.
XX
PN US6562958-B1.
XX
PD 13-MAY-2003.
XX
PF 04-JUN-1999; 99US-00328352.
XX
PR 09-JUN-1998; 98US-0088701P.
XX
PA (GENO-) GENOME THERAPEUTICS CORP.
XX
PI Breton G, Bush D;
XX
DR WPI; 2003-576092/54.
DR N-PSDB; ADA30178.
XX
PT New Acinetobacter baumannii proteins and nucleic acids, useful as reagents
PT for diagnosing a bacterial disease, as components of antibacterial
PT vaccines, as targets for antibacterial drugs, or as biocontrol agents for
PT plants.
XX
PS Example; SEQ ID NO 5591; 328pp; English.
XX
CC The invention relates to isolated Acinetobacter baumannii nucleic acids.
CC The A. baumannii nucleic acids and polypeptides are useful as reagents
CC for diagnosing a bacterial disease, as components of antibacterial
CC vaccines, as targets for antibacterial drugs, to detect the presence of
CC A. baumannii and other Acinetobacter species in a sample, in screening
CC compounds for the ability to interfere with the A. baumannii life cycle
CC or to inhibit A. baumannii infection, and as biocontrol agents for
CC plants. The present sequence represents the amino acid sequence of an A.
CC baumannii protein.
XX
SQ Sequence 275 AA;

Alignment Scores:
Pred. No.: 1.38e-26 Length: 275
Score: 478.50 Matches: 100
Percent Similarity: 61.9% Conservative: 51
Best Local Similarity: 41.0% Mismatches: 78
Query Match: 16.8% Indels: 15
DB: 6 Gaps: 5

US-10-665-990A-13 (1-1561) x ADA34304 (1-275)
QY 866 CGCATCGACTGCGCAGAGCGTCCAAACCGCTGTACGCGACACCCCTGCAAAAGAGCTC 925
Db ::
35 LysPheAspTrp-----ValLysAlaGluValValLysAspSerProAspLysIleArg 52
QY 926 GACCGCGACCGCGCAACCCCGGATTCCCGGAGGCTGCAAGACGCGCTCAAAACGCGCC 985
Db ::
53 SerLysAlaLysLysGluGluHisLeuAsnPheGlnLeuIleAsnHisLeuGluLysPro 72
QY 986 GAAAAACCGCTATCTCTGTTTCCACCTTATTCGTCCTACAAAATCCGCGACACGCA 1045
Db ::
73 GluSerAsnValAspLeuIleSerAlaTyrPheIleProGluLysGlnGlyAlaLysIle 92
QY 1046 CTGCGAAAACCTGTCGACGAGCGCATACGTTACCGTCTCTGACCAACTCGCTACAGCG 1105
Db ::
93 LeuSerThrLeuAlaLysGluGlyValGluValArgValLeuThrAsnSerPheLysAla 112
QY 1106 ACCGACGTTGCGCGCTCCATTCCGGCTACGTCAAATACCGAAAACCGCTGCTCAAAGCC 1165
Db ::
113 AsnAspValAlaValHisAlaPheTyrGlyLysTyrArgLysGluLeuLysAsn 132
QY 1166 GCATCAAACTCTAGGAGCTGCAACCC-----AACCAT 1198
Db ::
133 GlyValGlnLeuTyrGluPheLeuProThrProAspLysArgAspLeuAsnLysAsnThr 152
QY 1199 GCGCTCCCGCCACAAAAGACAAA---GGCTGACCGCGCAGCTCCGTAACACGCTGCAT 1255
Db ::
153 AspGluLeuAlaThrLysAlaLysValAsnMetLysGlyLeuSerArgSerSerLeuHis 172
QY 1256 GCCAAAACCTTCATTGTGGACGGCAACGCACTTTCATCGGCTATTCAACCTCGACGCC 1315
Db ::
173 ThrLysLeuMetAlaLeuAsp---GluGlnValPheIleGlySerPheAsnPheAspPro 191
QY 1316 CGTTCGCGACGCTCAATACGAAATGGCGTCGTCATCGAAAGCCCAAAATCCAGAA 1375
Db ::
192 ArgSerAlaTyrLeuAsnThrGluIleGlyValIleLeuAspSerProSerLeuAlaLys 211
QY 1376 CAGATGAGCGCACCTCGCCGATACACACCCGCAATACGCTACCGGTTACCTCGAC 1435
Db ::
212 ThrIleHisHisThrMetAspGluAsnLysTyrAlaTyrLysLysLysLeuAsp 231
QY 1436 AAACACACCGCTGCAATGGCAGAT-----CCGCGCACCCGAAAAACCTACCGGAC 1489
Db ::
232 ProAsnAsnHisIleTyrTrpGlnGlnGluThrProLysGlyProValIleTyrLysLys 251
QY 1490 GAAACCGGAAGCCAAACTTTGGAAACGCGATCGCGCAAAATCTCTATCTGCTGCCCATC 1549
Db ::
252 GluProGluMetLysTrpTrpGlnLysAlaGlyMetLysLeuLeuSerTrpLeuProLeu 271
QY 1550 GAAGGTTTATTA 1561
Db ::
272 GluGlyPheMet 275

RESULT 13
ABM88218
ID ABM88218 standard; protein; 2519 AA.
XX
AC ABM88218;
XX
DT 02-JUN-2005 (first entry)
XX
DE Rice abiotic stress responsive polypeptide SEQ ID NO:6464.
XX
KW abiotic stress tolerance; transgenic plant; plant; cereal; agriculture.
XX
OS Oryza sativa.
XX
PN WO2003008540-A2.
XX
PD 30-JAN-2003.
XX

```

21-JUN-2002; 2002WO-US019668.
22-JUN-2001; 2001US-0300112P.
24-AUG-2001; 2001US-0314662P.
26-SEP-2001; 2001US-0325277P.
21-NOV-2001; 2001US-0332132P.
XX
XX (SYGN ) SYNGENTA PARTICIPATIONS AG.
XX
XX Krepis J, Briggs SP, Cooper B, Glazebrook J, Goff SA, Katagiri F;
XX Moughamer T, Provart N, Ricke D, Zhu T;
XX
XX WPI; 2003-248011/24.
XX
XX New stress-responsive nucleic acid, useful for altering the
XX responsiveness of a plant, e.g. cereal, to an abiotic stress such as cold
XX stress, salt stress or osmotic stress.
XX
XX Claim 1; SEQ ID NO 6464; 89pp; English.
XX
XX The invention relates to novel abiotic stress responsive polynucleotides
XX and polypeptides. Also disclosed are vectors, expression cassettes, host
XX cells, and plants containing such polynucleotides. Also disclosed are
XX methods for using the polynucleotides and polypeptides to alter the
XX responsiveness of a plant to abiotic stress. The invention is useful in
XX agriculture. The nucleic acid is useful for determining whether a test
XX plant has been exposed to an abiotic stress condition. It is also useful
XX for selecting an agent that alters abiotic stress regulated
XX polynucleotide expression in a plant cell, and to identify a homolog or
XX ortholog to an abiotic stress responsive polynucleotide. The nucleic acid
XX molecule and the polypeptide encoded by it are useful in altering the
XX responsiveness of a plant to an abiotic stress, such as cold stress, salt
XX stress, osmotic stress or any of their combinations. The present sequence
XX is used in the exemplification of the invention
XX
XX Sequence 2519 AA;
XX
Alignment Scores:
Pred. No.: 4,05e-22 Length: 2519
Score: 421.00 Matches: 150
Percent Similarity: 41.9% Conservative: 76
Best Local Similarity: 27.8% Mismatches: 190
Query Match: 14.8% Indels: 123
DB: 7 Gaps: 18
XX
US-10-665-990A-13 (1-1561) x ABM88218 (1-2519)
QY 80 TGTTCAATGTTCTTCATGGTTGCCCCCATCGAAGACGG-----ACGGAAAGC----- 127
Db 377 CysAlaGlnAspSerProLeuArgAlaLeuAlaGlyArgAlaPheThrGluSerThrPro 396
QY 128 -----CGTCATTTCAACTCTCCAAACCTGTCTCTCTGGCAACATCTCTGCAATC 178
Db 397 GluGlnProArgHisTyrAlaThr----- 404
QY 179 CGGCACACCCCTCATAACAAACGGGCTATCCGACATCTACCTGCTCGACGACCCCAACGAA 238
Db 405 -----LeuLeuaspGluGlyGluGly 411
QY 239 GGCCTTGCGCGCGCGCGCCCTTATCGAATCTCGCGAAACACAGCTTCGATTTGCAATAC 298
Db 412 AlaLeuValAlaArgLeuAsnLeuArgSerAlaThrArgSerIleAspLeuGlnThr 431
QY 299 TACATTGGCGCAACGACATTTCCGGCAGGCTGTGTTCAACCTCATGTACTTGGCGCA 358
Db 432 TyrIlePheAspLeuAspSerAlaArgLeuValIleAspGluLeuLeuAlaIleSer 451
QY 359 GAACGCGCGCTCGCGGTACGCTGTGTTGACGACCAACACACGCGCGGTTGGACGAT 418
Db 452 ArgArgGlyValLysValArgValIleAspGlnLeuSerAlaIleSerAspLeuGln 471
QY 419 CTCCTGCTCGCCTCGAC---AGGCATCCCAATATCGAAGTGCCTGCGCTGTTCACCC--- 472

```

```
QY 1367 ATCGAGAACACAGATGAGCGCACCTCGCCGATACACACCCGAATACGCCTACCGCGTT 1426
Db |||||:::
837 PheAlaGluGlnLeu-----AlaGluSerIle 845
QY 1427 ACCCTCGACAAACACAAACCGCTGCAATGCGACGATCCGCGCACCCGAAACCT 1481
Db |||||:::
846 LeuArgAspIleHisProGlnAsnSerTrp-ThrValAlaProArgAlaLysPro 863

RESULT 14
ADA36398
ID ADA36398 standard; protein; 321 AA.
AC ADA36398;
XX
XX
XX 20-NOV-2003 (first entry)
XX Acinetobacter baumannii protein #3559.
XX Acinetobacter baumannii; bacterial disease; antibacterial; vaccine;
KW plant biocontrol agent.
XX Acinetobacter baumannii.
XX OS Acinetobacter baumannii.
XX PN US6562958-B1.
XX US6562958-B1.
XX PD 13-MAY-2003.
XX PF 04-JUN-1999; 99US-00328352.
XX PR 09-JUN-1998; 98US-0088701P.
XX PA (GENO-) GENOME THERAPEUTICS CORP.
XX PI Breton G, Bush D;
XX WPI: 2003-576092/54.
DR N-PSDB; ADA32272.
XX
XX New Acinetobacter baumannii proteins and nucleic acids, useful as reagents
PT for diagnosing a bacterial disease, as components of antibacterial
PT vaccines, as targets for antibacterial drugs, or as biocontrol agents for
PT plants.
XX Example; SEQ ID NO 7685; 328pp; English.
XX
XX The invention relates to isolated Acinetobacter baumannii nucleic acids.
CC The A. baumannii nucleic acids and polypeptides are useful as reagents
CC for diagnosing a bacterial disease, as components of antibacterial
CC vaccines, as targets for antibacterial drugs, to detect the presence of
CC A. baumannii and other Acinetobacter species in a sample, in screening
CC compounds for the ability to interfere with the A. baumannii life cycle
CC or to inhibit A. baumannii infection, and as biocontrol agents for
CC plants. The present sequence represents the amino acid sequence of an A.
CC baumannii protein.
XX
SQ Sequence 321 AA;

Alignment Scores:
Pred. No.: 1.88e-21 Length: 321
Score: 409.00 Matches: 100
Percent Similarity: 50.0% Conservative: 44
Best Local Similarity: 34.7% Mismatches: 110
Query Match: 14.3% Indels: 34
DB: 6 Gaps: 6

US-10-665-990A-13 (1-1561) x ADA36398 (1-321)

QY 74 CTTCTCTGTTCACTGTTCTTCATGGTTG-----CCCCCACTGGAGAACGGACGGAACG 127
Db ::::: |||||:::
44 IleValIleLysThrAsnAsnTrpLeuAsnAspGluAlaAlaGluGluLysThrGln--- 62

QY 128 CGTCATTCAATACTTCCAAACCTGTCTCTCTGGACAACTCTCTGCAATCTGCAATTCGGCACACC 187
```

```
Db 62 -----
QY 188 CTTATAACAAGCGCTATCCGACATCTACCTGCTCGACGACCCCCACGAAAGCCCTTGCC 247
Db |||||:::
63 -----GlnGlyLeuThrAlaPheValAlaLeuAspAlaPheMetSerIleAla 79
QY 248 GCCCGCGCGCCCTTATCGAATCTCCGAAACACAGCCTCGATTTCGAATCTACATTTGG 307
Db |||||:::
80 SerArgIleTyrLeuIleArgAsnAlaLysGluThrIleAspLeuGlnTyrTyrIleTrp 99
QY 308 CGCAACGACATTTCCGCGAGCGTCTGTTCACCTCATGTACCTTGCCTGCGCAGAACGCGC 367
Db |||||:::
100 ThrAsnAspPheValGlyAsnLeuIleHisGluLeuLeuLysAlaAlaAspArgGly 119
QY 368 GTGCGCGTACGCTGCTGTGGACGACAAACACGCGCGGTGGAGCATCTCTCTGCTC 427
Db |||||:::
120 IleLysValArgLeuLeuIleAspAspGlnAsnGlyIleLysLeuAspGlyIleLeuArg 139
QY 428 GCCCTCGACAGCCATCCCAATATCGAAGTGGCGCTGTTCAACCCCTTCGTCTCTAGCAA 487
Db |||||:::
140 SerLeuLeuGlnHisThrAsnPheGluIleArgLeuPheAsnProTyrLysPheArgTyr 159
QY 488 TGGCGCGCACTCGGCTACCTGACCGACTTCCCGCCCTCAACCGCGCATGCACAAACAA 547
Db |||||:::
160 LeuArgIlePheAspTyrLeuPheArgPheLysLysValAsnHisArgMethHisAsnLys 179
QY 548 TCCTTTACCGCGCAACCGCGCCACCATATCTCGCGGACGCAATATCGGCGACGAATAC 607
Db |||||:::
180 LeuIleIleAlaAspAlaSerIleAlaValThrGlyGlyArgAsnIleSerSerGluTyr 199
QY 608 TTCAAGTCCGTGAGGACACCGTTTTCGCGACCTGGACATCCTCGCACCGGACGCTC 667
Db |||||:::
200 PheGluAlaSerSerLysPheGlnPheThrAspMetAspIleLeuPheTyrGlyHisAla 219
QY 668 GTGCGCGAAGTATCGACGACTTCGACCGCTACTGGCAAGCATTCGCGCCCAACACGCG 727
Db |||||:::
220 ValArgHisAlaGlnAlaValPheThrAspPheTrpGluSerThrLeuSerValAsnAla 239
QY 728 AGCGCATCATCCGCGAGCGCAACATCGGCAAG---GGTCTTCAAGCATCTGGATACAAAC 784
Db |||||:::
240 ThrGluIleIle-----GlyThrCysAlaGluHisHisLeuLysAlaLeuArgGluHis 257
QY 785 GACGAAACA-----TCCAGACACGCGTCTCGCTACCGCGAAACCGTCGAAACAG 835
Db |||||:::
258 TyrGluGlnLeuHisHisGluAspHisSerLeuThrGluAspLysLeuTyrAspAlaGln 277
QY 836 TCGCGCCCTCTACCAAAAAATACAGACGGGACGCGCATCGATCGAGAGCGTCCAAACCCGC 895
Db |||||:::
278 SerTyrLeuLysGluLeuLeuGluHisAsnProIleGlnTrpSer-----LysAlaHis 295
QY 896 CTGATCAGGACACCCCTGCAAAA 919
Db |||||:::
296 PheValAlaAspSerProLysLys 303

RESULT 15
AEA04533
ID AEA04533 standard; protein; 5178 AA.
XX
XX AEA04533;
XX
XX 28-JUL-2005 (first entry)
XX
XX Human protein from gene under-expressed in cancer, MUC2.
XX Tumor marker; colon tumor; cancer; cytostatic; neoplaasm; diagnostic;
XX microarray; drug screening.
XX
XX Homo sapiens.
XX
XX WO2005044990-A2.
XX
XX 19-MAY-2005.
```


XX 01-NOV-2004; 2004WO-US036404.
 XX PF
 XX PR
 XX 04-NOV-2003; 2003US-00700439.
 XX PA (FARB) BAYER HEALTHCARE LLC.
 XX PA (MAYO-) MAYO FOUND MEDICAL EDUCATION & RES.
 XX PI Burgess C, Myerow S, Thiagalingam A, Maimonis P, Molino G;
 XX PI Burgart L, Boardman LA, Thibodeau S, Lewis M;
 XX DR WPI; 2005-372198/38.
 XX DR N-PSDB; AEA04440.
 XX DR REFSEQ; NP_002448.

PT Detecting expression of one or more nucleic acid sequences in biological
 PT sample, useful for detecting cancer, comprises detecting a change in the
 PT expression level of one or more nucleic acid sequences relative to a
 PT control expression level.

XX Claim 20; SEQ ID NO 178; 256pp; English.

XX The invention relates to detecting differential expression of one or more
 CC nucleic acid sequences (appearing as AEA04356-AEA04448 in a biological
 CC sample comprising obtaining the sample from a subject, and detecting a
 CC change in the expression level of one or more nucleic acid sequences
 CC relative to a control expression level of the nucleic acid sequences,
 CC new. Also included are detecting cancer (or a pre-malignant condition
 CC thereof) in a subject (comprising comparing the expression level of one
 CC or more nucleic acid sequences in a biological sample from the subject
 CC with a control expression level of the nucleic acid sequences, where a
 CC change of at least two-fold in the expression level of the nucleic acid
 CC sequences is indicative of cancer or pre-malignant condition), monitoring
 CC the onset (or progression, or regression) of cancer (or a pre-malignant
 CC condition) in a subject (by detecting in a biological sample of the
 CC subject at a first point in time the expression of one or more nucleic
 CC acid sequences, repeating the first step at a subsequent point in time
 CC and comparing the expression level detected, where a change in the
 CC expression level is indicative of progression of cancer or its pre-
 CC malignant condition in the subject), determining prognosis for cancer or
 CC its pre-malignant condition in a subject (comprising detecting in a
 CC biological sample of the subject, the expression level of one or more
 CC nucleic acid sequences, comparing the expression level with a reference
 CC expression level of the nucleic acid sequences and evaluating the
 CC prognosis of the subject based on the comparison), determining the
 CC efficacy of a test compound for inhibiting cancer in a subject,
 CC determining the efficacy of a therapy for inhibiting cancer in a subject,
 CC selecting a composition for inhibiting cancer in a subject, inhibiting
 CC cancer in a subject, a polypeptide encoded by the nucleic acids above
 CC (appearing as AEA04449-AEA04541), an antibody that specifically binds to
 CC the polypeptide sequence, and detecting in a biological sample the
 CC presence of a polypeptide. The method is useful for detecting in a
 CC differential expression of one or more nucleic acid sequences in a
 CC biological sample, which is useful for detecting cancer (especially colon
 CC cancer), monitoring the onset, progression, or regression of cancer or a
 CC pre-malignant condition, or determining prognosis for cancer or its pre-
 CC malignant condition in a subject, or for determining the efficacy of a
 CC test compound for inhibiting cancer in a subject. The compound is useful
 CC for inhibiting cancer in a subject. The antibodies may also be used to
 CC treat cancer. The present sequence is a protein from a human gene under-
 CC expressed in cancer samples.

XX Sequence 5178 AA;

Alignment Scores:

Pred. No.:	8,08e-16	Length:	5178
Score:	336.50	Matches:	141
Percent Similarity:	33.6%	Conservative:	55
Best Local Similarity:	24.1%	Mismatches:	198
Query Match:	11.8%	Indels:	190
DB:	9	Gaps:	22

US-10-665-990A-13 (1-1561) x AEA04533 (1-5178)

QY	119	ACGGAAGCGCTGCTTTTCAAT-----ACTTCCAAACCTGTCTCTCTGGACAC	166
Db	1248	ThrValGluLysHisPheAsnIleCysSerIleThrThrArgPro-SerThrLeuThrTh	1267
QY	167	ATCTTGCAAAATCCGGCACACCCCTCATAAACACGGGTATCCGACATCTACTGCTCGAC	226
Db	1267	rPheThrThrIleThrLeuThrThrProThrProThrSerPheThrThrThrThrTh	1287
QY	227	GACCCCAACGAGCGCTTGGCCGCCGCCCTTATCGAATCTGCCGAACACAGCCTC	286
Db	1287	rThrProThr-----SerSerTh-----	1293
QY	287	GATTTGCAATATACATTGGCGCAACGACATTTCCGGCAGCGCTGCTTCAACCTCA--	344
Db	1293	rValLeuSerThrThr-----ProLysLeuCysCysLeuThrProSerAs	1307
QY	344	-----	344
Db	1307	pTrpIleAsnGluAspHisProSerSerGlySerAspAspGlyAspArgGluProPheAs	1327
QY	345	-----TGTAACCTTGGCGAGAACGGCGCGTGGCTACGCC-----	380
Db	1327	pGlyValCysGlyAlaProGluAspIleGluCysArgSerValLysAspProHisLeuSe	1347
QY	380	-----	380
Db	1347	rLeuGluGlnHisGlyGlnLysValGlnCysAspValSerValGlyPheIleCysLysAs	1367
QY	380	-----	380
Db	1367	nGluAspGlnPheGlyAsnGlyProPheGlyLeuCysTyrAspTyrLysIleArgValAs	1387
QY	381	TGCTGTGG-----	400
Db	1387	nCysCysTrpProMetAspLysCysIleThrThrProSerProThrThrThrProSe	1407
QY	401	ACGGCGGGTGGACGATCTCTGCTCGCTCGCTCGACGCCATCCCAATATCGAAGTGGC	460
Db	1407	rProProProThrThrThrThrLeuProProThrThrPro-----	1422
QY	461	CTGTTCACCCCTTCGTCCTACGCAAAATGGCGCACCTCGGCTACCTGACGACTTCC	520
Db	1423	-----SerProProThrThrThrThrProProProThrThrThrProSerProPr	1441
QY	521	CGCTCAACCGCGCATGCACACAAATCTTTACCGCCGACACCGCGCCACCATCTC	580
Db	1441	oIleThrThr-----ThrThrThrProLeuProThrThrThrProSerPro-----	1456
QY	581	GGCGGCGCAATATCGCGCAGCAATCTTCAAAGTCGGTAGGACACCGTTTCGCGCAC	640
Db	1457	-----ProIleSerThrThrThrThrProProProThrThrThrProSerProProTh	1474
QY	641	CTGGACATCTCGCCACCGCAGCGTGTGGGGAAGTATCGCACGACTTCGAC-----	695
Db	1474	rThrThrProSerProProThrThrThrProSerProProThrThrThrThrThrPr	1494
QY	696	-----GCTACTGGGCAACCCATTCCGCCCAACACGCCGCGCATCTCCGCGCGGC	748
Db	1494	oProProThrThrThrProSerProProMetThrThrProIleThrProProAlaSerTh	1514
QY	749	AACATCGGCAAGGTCTTCAAGCACTCGGATACAAACGACCAACATCCAGACACCGCTC	808
Db	1514	rThrThrLeu-----ProProThrThrThrProSerProProThrThrThrTh	1529
QY	809	CTGGCTACCGGAAACCGTCGAAC-----AGTCGCCCTCTACCAAAAAATACAGACG	862
Db	1529	rThrThrThrProProProThrThrThrProSerProProThrThr-----	1544
QY	863	GGAGGATCGACTGGCAGACGCTCCAAACCCGCTGATCAGCGACCCCTCGAAAAAGGA	922
Db	1545	-----ThrProIleThrProPr-----	1550

```
QY 923 CTCGACCGCGCGCGGCAAAACCGCGGATTCGGGGAGGCTGCAAGACCGCGTCAAAACAG 982
Db : : : : : |||
1550 oThrSerThrThrThrThrThrThrThrThrThrThrThrThrThrThrThrThrThrThrThr
QY 983 CCGCAAAAAGGCGTCTATCTGTTTCACCTATTTCGTCCTTACAAAATCCGGGCACAGAC 1042
Db : : : : : |||
1563 rProProProThrThrThrThrThrThrThrThrThrThrThrThrThrThrThrThrThrThrThr
QY 1043 GCACTGGCAAAACTGGTCAGACGGCATAGACGTTACCGTCTCTGACCA-----AC 1093
Db : : : : : |||
1583 r-----ThrThrProSerProProThrThrThrThrThrThrThrThrThrThrThrThrThr
QY 1094 TCGCTACAGCGCACCGAGTTGCCCGCTCCGCTCATTGCGGCTAGCTCAAAAT----- 1142
Db : : : : : |||
1593 rThrThrProProProThrThrThrThrThrThrThrThrThrThrThrThrThrThrThrThrThr
QY 1143 -----ACGAAAAACCGCTGCTCAAAAGCGGCATCAAACTC 1177
Db : : : : : |||
1613 oProThrThrThrThrThrThrThrThrThrThrThrThrThrThrThrThrThrThrThrThrThr
QY 1178 TACGAGCTGCAACCAACGATCCGTCGCCCAAAAGACAAAGGCTGACCGGCAGC 1237
Db : : : : : |||
1633 rThrLeuProProThrThrThrProSerProProProThrThrThrThrThrThrThrThrThrThr
QY 1238 TCGCTAACCGAGCTGCATGCCAAACCTTCATTGTGGACGGCAACGCGATCTTCATCGCG 1297
Db : : : : : |||
1650 rProProProThrThrThrThrThrThrThrThrThrThrThrThrThrThrThrThrThrThrThr
QY 1298 TCATTCAACCTCGACCCCGTTCGCCACGCTCAATACCGAAATGGCGTCTCATCGAA 1357
Db : : : : : |||
1669 eThrThrThrThrThrThrThrThrThrThrThrThrThrThrThrThrThrThrThrThrThrThr
QY 1358 AGCCCCAAATCGAGACAGATGAGCGCACCTCGCGATACACACCCGGAATACGCC 1417
Db : : : : : |||
1689 oSerProProThr-----ThrThrMetThrThrThrProSerProThrThrThrThrThrThrThr
QY 1418 TACCGGTTACCTCGACAAACACACACCGCTGCAATGGCAGATCCCGCCACCCGAAAA 1477
Db : : : : : |||
1708 oile-----ThrThrThrThrThrThrThrThrThrThrThrThrThrThrThrThrThrThrThr
QY 1478 ACCTACCCGACGAAACCGCCAAACCTTTGGAAACGCGATCGCGCAAAAATCTCTATCC 1537
Db : : : : : |||
1725 oThrThrMetThrThrThrProSerProThrThrThrThrThrThrThrThrThrThrThrThrThr
QY 1538 CTGCTGCCA 1547
Db : : : : : |||
1745 rThrLeuPro 1748
RESULT 16
AAM24516
ID AAM24516 standard; protein; 5179 AA.
XX AAM24516;
AC AAM24516;
XX
XX 12-OCT-2001 (first entry)
XX
XX C899P predicted amino acid sequence.
DE Human; immunotherapy; diagnosis; colon cancer; colon tumour; immunogenic;
XX gene therapy; vaccine; colonic cancer.
XX
XX Homo sapiens.
XX
XX WO200149716-A2.
XX
XX 12-JUL-2001.
XX
XX 29-DEC-2000; 2000WO-US035596.
XX
XX 30-DEC-1999; 99US-00476296.
PR 10-JAN-2000; 2000US-00480321.
PR 15-FEB-2000; 2000US-00504629.
PR 06-MAR-2000; 2000US-00519444.
```

```
PR 19-MAY-2000; 2000US-00575251.
PR 29-JUN-2000; 2000US-00609448.
PR 28-AUG-2000; 2000US-00649811.
XX (CORI-) CORIXA CORP.
XX
XX PA
XX Xu J, Lodes MJ, Secrist H, Benson DR, Meagher MJ, Stolk JA;
PI King GE, Wang T, Jiang Y;
XX WPI; 2001-441847/47.
XX
XX Colon tumor associated proteins and nucleic acids useful for the
PT prevention, diagnosis and treatment of colonic cancer.
XX
XX Claim 2; Page 446-462; 472pp; English.
XX
XX The present invention describes colon tumour associated proteins (I) and
CC the polynucleotides (II) that encode them. (I) have cytostatic activity,
CC (I) and (II) can be used in gene therapy and vaccine production. (I) and
CC (II) may be used in the prevention, diagnosis and treatment of diseases
CC associated with inappropriate colon tumour associated protein (TCAP)
CC expression, such as colonic cancer. For example, (I) and (II) may be used
CC to treat disorders associated with decreased expression by rectifying
CC mutations or deletions in a patient's genome that affect the activity of
CC TCAPs by expressing inactive proteins or to supplement the patients own
CC production of them. Additionally, (II) may be used to produce the TCAP
CC proteins, by inserting the nucleic acids into a host cell culturing the
CC cell to express the protein. (II) and its complementary sequences may
CC also be used as DNA probes in diagnostic polymerase chain reaction (PCR)
CC and hybridisation assays to detect and quantitate the presence of similar
CC nucleic acids in samples, and therefore which patients may be in need of
CC restorative therapy. (I) may also be used as antigens in the production
CC of antibodies against TCAPs and in assays to identify modulators of TCAP
CC expression and activity. Anti-(I) antibodies and antagonists may also be
CC used to down regulate TCAP expression and activity. The anti-(I)
CC antibodies may also be used as diagnostic agents for detecting the
CC presence of TCAPs in samples (e.g. by enzyme linked immunosorbant assay
CC (ELISA)). AAM248460 to AAM29512 and AAM24494 to AAM24523 represent
CC nucleotide and amino acid sequences given in the exemplification of the
CC present invention
XX
XX SQ Sequence 5179 AA;
Alignment Scores:
Pred. No.: 8.08e-16 Length: 5179
Score: 336.50 Matches: 141
Percent Similarity: 33.6% Conservative: 55
Best Local Similarity: 24.1% Mismatches: 198
Query Match: 11.8% Indels: 190
DB: 4 Gaps: 22
US-10-665-990A-13 (1-1561) x AAM24516 (1-5179)
QY 119 ACGGAAAGCCGTCATTTCAT------ACTTCCAAACCTGCTCTCTCGGACAAAC 166
Db : : : : : |||
1248 ThrValGluLysHisPheAsnIleCysSerIleThrThrArgPro-SerThrLeuThrTh 1267
QY 167 ATCTCGCAAAATCCGGCACACCCCTCATAAACGGGCTATCCGACATCTACTGCTCGAC 226
Db : : : : : |||
1267 rPheThrThrIleThrLeuProThrThrProThrSerPheThrThrThrThrThrThrThr 1287
QY 227 GACCCCGACGAGCCCTTCCGCCCGCCGCCCTTATCGAATCTGCCGACACAGCCTC 286
Db : : : : : |||
1287 rThrProThr-----SerSerTh 1293
QY 287 GATTTCGAATACTACATTGTGGCGCAACGACATTTCCGAGAGCTGCTGTTCACCTCA-- 344
Db : : : : : |||
1293 rValLeuSerThrThr-----ProlLysLeuCysCysLeuTrpSerAs 1307
QY 344 ----- 344
Db 1307 pTrpIleAsnGluAspHisProSerSerGlySerAspGlyAspArgGluProPheAs 1327
```


CC development of cancer in a patient. (1) have immunostimulant and
 CC cytostatic activities and can be used in vaccines. AB232646 to AB233725
 CC and ABP55343 to ABP55391 represent human colon cancer/tumour related
 CC sequences used in the exemplification of the present invention

XX Sequence 5179 AA;

Alignment Scores:

Pred. No.: 8.08e-16 Length: 5179
 Score: 336.50 Matches: 141
 Percent Similarity: 33.6% Conservative: 55
 Best Local Similarity: 24.1% Mismatches: 198
 Query Match: 11.8% Indels: 190
 DB: 6 Gaps: 22

US-10-665-990A-13 (1-1561) x ABP55365 (1-5179)

```

QY 119 ACGGAAAGCGCTATTCTCAAT-----ACTTCCAAACCTGTCTCTCTGGACAAC 166
DB 1248 ThrValGluLysHisPheAsnIleCysSerIleThrThrArgPro-SerThrLeuThrTh 1267
QY 167 ATCTCGCAATCCGGCACACCCCTCATAAACACGGCTATCCGACATCTACTGCTCGAC 226
DB 1267 rPheThrThrIleThrLeuProThrThrProThrSerPheThrThrThrThrTh 1287
QY 227 GACCCCAAGACCGCTTGGCGCCCGCGCCCTTATCGAATTCGCGAACAACAGCCTC 286
DB 1287 rThrProThr-----SerSerTh 1293
QY 287 GATTGCAATACTACATTGGCGCAACGACATTTCCGGCAGGCTGCTGTCAACCTCA-- 344
DB 1293 rValLeuSerThrThr-----ProLysLeuCysCysLeuTrpSerAs 1307
QY 344 ----- 344
DB 1307 TrpIleAsnGluAspHisProSerSerGlySerAspGlyAspArgGluProPheAs 1327
QY 345 -----TGTACCTTGGCGAGAACGGCGGTGGCGGTACGGC----- 380
DB 1327 pGlyValCysGlyAlaProGluAspIleGluCysArgSerValLysAspProHisLeuSe 1347
QY 380 ----- 380
DB 1347 rLeuGluGlnHisGlyGlnLysValGlnCysAspValSerValGlyPheIleCysLysAs 1367
QY 380 ----- 380
DB 1367 nGluAspGlnPheGlyAsnGlyProPheGlyLeuCysTyraSpTyrlLysIleArgValAs 1387
QY 381 -TGCTGTTGG-----ACGACAAACAAC 400
DB 1387 nCysCysTrpProMetAspLysCysIleThrThrProSerProProThrThrProSe 1407
QY 401 ACGCGCGGTTGGAGGATCTCTGCTCGCCCTCGACAGCATCCCAATATCGAAGTGGC 460
DB 1407 rProProProThrThrThrThrThrLeuProProThrThrPro----- 1422
QY 461 CTGTTCAACCCCTCGTCTACGCAAAATGGCGGCACTCGGCTACTGACCGACTTCCC 520
DB 1423 ---SerProProThrThrThrThrThrProProProThrThrThrProSerProPr 1441
QY 521 CGCCTCAACCGCGGATGACAAACAAATCTTTACCGCGGCAACCGCGGCACCACTACTC 580
DB 1441 oIleThrThr-----ThrThrProLeuProThrThrThrProSerPro----- 1456
QY 581 GCGCAGCAGCAATATCGGACCAATACTTCAAAGTCGGTGAGGACACACGCTTTTCGCCGAC 640
DB 1457 -----ProIleSerThrThrThrThrProProProThrThrThrProSerProProTh 1474
QY 641 CTGGACATCTCGCACCGCGGACGCTCGTGGCGAGTATCGCAGACTTCGACC----- 695
DB 1474 rThrThrProSerProProThrThrThrProSerProProThrThrThrThrThrPr 1494

```

```

QY 696 -----GCTACTGGCAAGCCATTTCGCCCAACAGCCGCGCATCATCCGCGAGCGC 748
DB 1494 oProProThrThrThrProSerProMetThrThrProIleThrProProAlaSerTh 1514
QY 749 AACATCGCAAGGCTCTTAAGCACTCGGATACAAACGACAAACATCCAGACACGCGCTC 808
DB 1514 rThrThrLeu-----ProProThrThrThrProSerProProThrThrTh 1529
QY 809 CTGCGCTACCGCGAAACCGTGAAC-----AGTCGCCCTCTACCAAAAAATACAGACG 862
DB 1529 rThrThrProProThrThrThrProSerProProThrThr----- 1544
QY 863 GGACGCTGACTGGCAGAGCGTCCAAACCGCCTGATCAGCGACACCCCTGCAAAAGGA 922
DB 1545 -----ThrProIleThrProPr 1550
QY 923 CTCGACCGCGACCGCGCAACCGCGATTTCGGGAGGCTGCAAGACGCGCTCAAAACAG 982
DB 1550 oThrSerThrThrThr-----LeuPro-----ProThrThrThrProSe 1563
QY 983 CCGGAAAAAGCGTCTATCTGGTTTCACCCCTATTTGTCCTTACAAAATCCGGCAGAC 1042
DB 1563 xProProProThrThrThrProProProThrThrThrProSerProProThrTh 1583
QY 1043 GCACCTGGCAAACTGGTGCAGGACGCGATACACCTTACCGTCTGACCA-----AC 1093
DB 1583 r-----ThrThrProSerProProThrIleThrTh 1593
QY 1094 TCGCTACAGGCGACGCGTTCGCGCTTCATTCGCGTACGTCAAT----- 1142
DB 1593 rThrThrProProThrThrThrProSerProProThrThrThrThrThrThrProPr 1613
QY 1143 -----ACGAAAAACGCTGTCTCAAAGCGCGCATCAAACTC 1177
DB 1613 oProThrThrThrProSerProThrThrThrProIleThrProThrSerThrTh 1633
QY 1178 TACGAGCTGCAACCAACCATGCGTCCCGCCGCAACAAAGACAAAGGCTGACCGGACG 1237
DB 1633 rThrLeuProProThrThrThrProSerProProProThrThr-----ThrThrTh 1650
QY 1238 TCCGTAAACCGCTGCGATGCCAAAACCTTCATTGTGGACGGCAACACGATCTTCATCGC 1297
DB 1650 rProProProThrThrThrProSerProProThr---ThrThrThrProSerProPr 1669
QY 1298 TCATTCAACCTGACGCCCGCTTCGCGACGGCTCAATACCGAAATGGGCGTCTATCGAA 1357
DB 1669 eThrThrThrThrProProProThrThrThrProSerSerProIleThrThrPr 1689
QY 1358 AGCCCAAAATCGCAGAACAGATGGAGCGCACCTCGCGCATACCAACCCGAAATACGCC 1417
DB 1689 oSerProProThr---ThrThrMetThrThrProSerProThrThrThrProSerSerPr 1708
QY 1418 TACCGCGTTACCTCGACAAACACACCGCTGCAATGGCAGATCCCGCACCCGCAAAA 1477
DB 1708 oIle-----ThrThrThrThrProSerSerThrThrThrThrProSerProPr 1725
QY 1478 ACCTACCGAAGCAACCGCAAGCAAACTTTGGAAACGCATCGCGGCAAAATCCTATCC 1537
DB 1725 oThrThrMetThrThrProSerProThrThrThrProSerProProThrThrMetTh 1745
QY 1538 CTGCTGCCCA 1547
DB 1745 rThrLeuPro 1748

```

RESULT 18

ABO07258
 ID ABO07258 standard; protein; 5179 AA.
 XX ABO07258;
 XX 13-AUG-2003 (first entry)
 XX Human p53 modifying protein, SEQ ID 218.

XX Human; p53 modifier; cytostatic; cancer; cytostatic; antiangiogenic;
 KW antiapoptotic; p53 pathway; breast cancer; colon cancer; kidney cancer;
 KW lung cancer; ovarian cancer; angiogenesis; cell cycle;
 KW apoptotic disorder; cell proliferation disorder.

XX Homo sapiens.

PN WO200299122-A1.

XX 12-DEC-2002.

PF 03-JUN-2002; 2002WO-US017382.

XX 05-JUN-2001; 2001US-0296076P.

PR 10-OCT-2001; 2001US-0328605P.

PR 15-FEB-2002; 2002US-0357253P.

XX (EXEL-) EXELIXIS INC.

XX Friedman L, Plowman GD, Belvin M, Francis-Lang H, Li D, Funke RP;
 PI N-PSDB; ACD13431.

DR WPI; 2003-156859/15.

XX Identifying modulators of the p53 pathway for use in treating apoptotic

PT or cell proliferation disorders, comprises screening for agents that

PT modulate activity of a human ortholog of genes that modify the p53

PT pathway in Drosophila.

XX Example 2; Page 623-639; 678pp; English.

XX The invention relates to identifying (M1) a candidate p53 pathway
 CC modulating agent, by contacting an assay system comprising a purified HM
 CC polypeptide (human orthologue of genes that modify the p53 pathway in
 CC Drosophila) or nucleic acid with a test agent under conditions, where but
 CC for the presence of the test agent, the system provides a reference
 CC activity, and detecting a test agent-biased activity of the assay system.
 CC Also included are modulating (M2) a p53 pathway of a cell (comprising
 CC contacting a cell defective in p53 function with a candidate modulator
 CC that specifically binds to a HM polypeptide comprising an HM amino acid
 CC sequence, where p53 function is restored; modulating (M3) a p53 pathway
 CC in a mammalian cell (comprising contacting the cell with an agent that
 CC specifically binds an HM polypeptide or nucleic acid) and diagnosing (M4)
 CC a disease in a patient (comprising: (a) obtaining a biological sample
 CC from the patient; (b) contacting the sample with a probe for HM
 CC expression; (c) comparing the results with a control; and (d) determining
 CC whether the comparison indicates a likelihood disease). (M1) is useful
 CC for identifying modulators of the p53 pathway. A probe for HM expression
 CC is useful for diagnosing breast, colon, kidney, lung and ovarian cancer,
 CC in a patient, where the cancer has greater than 25 % expression level.
 CC Modulators identified by (M1) are useful in a variety of diagnostic and
 CC therapeutic applications, where disease or disorder prognosis is related
 CC to defects in the p53 pathway, such as, angiogenesis, apoptotic or cell
 CC proliferation disorders (e.g. cancer). Another two new methods (M2 and
 CC M3) are useful for modulating the p53 pathway of a cell, thus restoring
 CC the p53 function of the cell, so that the cell undergoes normal
 CC proliferation or progression through the cell cycle. (M2) and (M3) are
 CC also useful for treating defects in the p53 pathway such as angiogenic,
 CC apoptotic or cell proliferation disorders. The present sequence
 CC represents a human p53 pathway modifying protein

XX SQ Sequence 5179 AA;

Alignment Scores:

Pred. No.:	8.08e-16	Length:	5179
Score:	336.50	Matches:	141
Percent Similarity:	33.6%	Conservative:	55
Best Local Similarity:	24.1%	Mismatches:	198
Query Match:	11.8%	Indels:	190
DB:	6	Gaps:	22

US-10-665-990A-13 (1-1561) x ABO07258 (1-5179)

QY	119	ACGGAAGCGCTCATTTTCAAT-----ACTTCCAAACCTGTCTCTCGACAC	166
Db	1248	ThrValGluLysHisPheAsnIleCysSerIleThrThrArgPro-SerThrLeuThrTh	1267
QY	167	ATCTGCAAAATCCGGCACACCCCTCATAAACACGGGTATCCGACATCTACCTGCTGCAC	226
Db	1267	rPheThrThrIleThrLeuProThrThrProThrSerPheThrThrThrThrThrTh	1287
QY	227	GACCCCGACGAAGCCCTTGGCGCGCGCCCTTATCGAATCTGCCGAACACAGCCTC	286
Db	1287	rThrProThr-----SerSerTh	1293
QY	287	GATTTCGAATACATTTGGCGCAACGACATTTCCGGCAGGCTGCTGTCAACCTCA--	344
Db	1293	rValLeuSerThrThr-----ProLysLeuCysCysLeuTrpSerAs	1307
QY	344	-----	344
Db	1307	pTrpIleAsnGluAspHisProSerSerGlySerAspGlyAspArgGluProPheAs	1327
QY	345	-----TGTAACCTTGGCGCAGACGGCGGTGGCGGTACGCC-----	380
Db	1327	pGlyValCysGlyAlaProGluAspIleGluCysArgSerValLysAspProHisLeuSe	1347
QY	380	-----	380
Db	1347	rLeuGluGlnHisGlyGlnLysValGlnCysAspValSerValGlyPheIleCysLysAs	1367
QY	380	-----	380
Db	1367	nGluAspGlnPheGlyAsnGlyProPheGlyLeuCysTyrAspTyrLysIleAlaValAs	1387
QY	381	-TGCTGTGG-----ACGACAAACAC	400
Db	1387	nCysCysTrpProMetAspLysCysIleThrThrProSerProThrThrThrProSe	1407
QY	401	ACGCGCGGTGGACGATCTCTCTCGCTCGCTCGACGACCATCCCAATATCGAAGTGGC	460
Db	1407	rProProThrThrThrThrThrLeuProThrThrPro-----	1422
QY	461	CTGTCAACCCCTTCTGCTACGCAAAATGGCGCGCTACCTGACGACTTCCCC	520
Db	1423	-----SerProProThrThrThrProProProThrThrThrProSerProPr	1441
QY	521	GGCTCAACCGCCGATGCAACAATACTTTTACCGCGCACACCGCGCCACCATATCTC	580
Db	1441	oIleThrThr-----ThrThrThrProLeuProThrThrThrProSerPro-----	1456
QY	581	GGCGGACGCAATATCGCGACGAATACTTCAAAGTCGGTGAGGACACCGTTTTCGCCAC	640
Db	1457	-----ProIleSerThrThrThrThrProProProThrThrThrProSerProProTh	1474
QY	641	CTGACATCTCTCGCACCGGAGCGTTCGTCGGGAGATATGCGACGACTTCGAC--	695
Db	1474	rThrThrProSerProProThrThrThrProSerProProThrThrThrThrThrPr	1494
QY	696	-----GCTACTGGGCAAGCCATTCCGCCCAACGCGCGCATCATCCGCGAGCGGC	748
Db	1494	oProProThrThrThrProSerProMetThrThrProIleThrProProAlaSerTh	1514
QY	749	AACATCGGAAGGGTCTTCAAGCACTCGGATACCAACGACGCAAAACATCCAGACGCGCTC	808
Db	1514	rThrThrLeu-----ProProThrThrThrThrProSerProProThrThrTh	1529
QY	809	CTGCGCTACCGCAAAACCGTCGAAC-----AGTCGCCCTCTTACCAAAAAAATACAGACG	862
Db	1529	rThrThrThrProProProThrThrThrProSerProProThrThr-----	1544
QY	863	GGACGATCGACTGGGAGAGCGGTCCAAACCGCGCTGATCAGCGCACCCCTCGCAAAAGGA	922
Db	1545	-----ThrProIleThrProPr	1550

QY 983 CCGGAAAAAGCGTCTATCTGTTTCACCCCTATTTCGTCCTACAAAATCCGGCAGAC 1042
 ||||| : : : : :
 Db 1563 rProProThrThrThrThrThrThrThrThrThrThrThrThrThrThrThrTh 1583
 ||||| : : : : :
 QY 1043 GCACCTGGCAAAACTGGTGCAGCAGCATAGACGTTACCGTCTCGACCA-----AC 1093
 ||||| : : : : :
 Db 1583 r-----ThrThrProSerProProThrThrThrThrThrThrThrThrThr 1593
 ||||| : : : : :
 QY 1094 TCGCTACAGGACCGACGCTGCCCGCTCCATTCGGCTACGTCAGTCAAAAT----- 1142
 ||||| : : : : :
 Db 1593 rThrThrProProThrThrThrThrThrThrThrThrThrThrThrThrThrThrThr 1613
 ||||| : : : : :
 QY 1143 -----ACGAAAACCGCTCTCAAGCCGGCATCAAACTC 1177
 ||||| : : : : :
 Db 1613 oProThrThrThrThrThrThrThrThrThrThrThrThrThrThrThrThrThrThr 1633
 ||||| : : : : :
 QY 1178 TAGGAGCTCAACCAACCATCCGTCGCCGCGCACAAAAGCAAGGCTGACCGGCAGC 1237
 ||||| : : : : :
 Db 1633 rThrLeuProProThrThrThrThrThrThrThrThrThrThrThrThrThrThrThr 1650
 ||||| : : : : :
 QY 1238 TCCGTAACGAGCTGCATGCCAAAACCTTCATTTGTGAGCGCAAAACGCATCTTCATCGGC 1297
 ||||| : : : : :
 Db 1650 rProProThrThrThrThrThrThrThrThrThrThrThrThrThrThrThrThrThr 1669
 ||||| : : : : :
 QY 1298 TCATTCAACCTCGACCCCGTTCCGACCGCTCAATACCGAATGGCGTCTCATCGAA 1357
 ||||| : : : : :
 Db 1669 eThrThrThrThrThrThrThrThrThrThrThrThrThrThrThrThrThrThrThr 1689
 ||||| : : : : :
 QY 1358 AGCCCAAAATCGCAGACAGATGAGCGCACCCCTCGCGATACACACCCGAAATACGCC 1417
 ||||| : : : : :
 Db 1689 oSerProProThrThrThrThrThrThrThrThrThrThrThrThrThrThrThrThrThr 1708
 ||||| : : : : :
 QY 1418 TACCGGTTTACCTCGACAAAACAAACCGCTTCATGCAATGCGACGATCCCGCCACCCGAAAA 1477
 ||||| : : : : :
 Db 1708 oile-----ThrThrThrThrThrThrThrThrThrThrThrThrThrThrThrThr 1725
 ||||| : : : : :
 QY 1478 ACTTACCGAAGCAGCCGAGCCAACTTTGGAACGATCGCGCAAAATCTCATCC 1537
 ||||| : : : : :
 Db 1725 oThrThrMetThrThrThrThrThrThrThrThrThrThrThrThrThrThrThrThrThr 1745
 ||||| : : : : :
 QY 1538 CTGCTGCCCA 1547
 ||||| : : : : :
 Db 1745 rThrLeuPro 1748
 ||||| : : : : :
 RESULT 22
 ID ADQ80379
 ADQ80379 standard; protein; 5179 AA.
 AC ADQ80379;
 XX
 DT 21-OCT-2004 (first entry)
 XX
 DE Intestinal/tracheal mucin 2 protein.
 XX
 KW cytostatic; epidermal growth factor receptor modulator; identification;
 KW therapeutic response; cancer; EGFR; biomarker.
 XX
 OS Homo sapiens.
 XX
 PN WO2004063709-A2.
 XX
 PD 29-JUL-2004.
 XX
 PF 08-JAN-2004; 2004WO-US000368.
 XX
 PR 08-JAN-2003; 2003US-0438735P.
 XX
 PA (BRIM) BRISTOL-MYERS SQUIBB CO.
 XX
 PI Amler LC, Januario T;
 XX
 DR WPI; 2004-544114/52.
 DR N-PSDB; ADQ80262.

XX Identifying a mammal that will respond therapeutically to a method of
 PT treating cancer comprises comparing the level of a biomarker in a mammal
 PT before and after exposure to an epidermal growth factor receptor (EGFR)
 PT modulator.
 XX
 PS Disclosure; SEQ ID NO 151; 520pp; English.
 XX
 CC The invention relates to a method of identifying a mammal that will
 CC respond therapeutically to a method of treating cancer by administering
 CC an epidermal growth factor receptor (EGFR) modulator by comparing the
 CC level of a biomarker in a mammal before and after exposure to an EGFR
 CC modulator. The method comprises: (a) measuring, in the mammal, the level
 CC of at least one biomarker identified in the specification; (b) exposing
 CC the mammal to the EGFR modulator; and (c) measuring in the mammal the
 CC level of the biomarker, where a difference in the level in step (c)
 CC compared to step (a) indicates that the mammal will respond
 CC therapeutically to the method of treating cancer. The method and
 CC biomarkers are useful for identifying a mammal that will respond
 CC therapeutically to a method of treating cancer by administering an
 CC epidermal growth factor receptor (EGFR) modulator. This sequence
 CC corresponds to one of the biomarkers whose levels of expression is
 CC measured in the method of the invention.
 XX
 SQ Sequence 5179 AA;
 8.08e-16 Length: 5179
 Pred. No.: 141
 Score: 336.50 Matches: 141
 Percent Similarity: 33.6% Conservative: 55
 Best Local Similarity: 24.1% Mismatches: 198
 Query Match: 11.8% Indels: 190
 DB: Gaps: 22
 US-10-665-990A-13 (1-1561) x ADQ80379 (1-5179)
 QY 119 ACGGAAAGCCGCTCATTTCAAT-----ACTTCCAAACCTGTCTCTCTGCGAAC 166
 ||||| : : : : :
 Db 1248 ThrValGluLysHisPheAsnIleCysSerIleThrThrArgPro-SerThrLeuThrTh 1267
 ||||| : : : : :
 QY 167 ATCTCTGCAATCCGGCACACCCCTCATAAACAGGGCTATCCGACATCTACCTGCTCGAC 226
 ||||| : : : : :
 Db 1267 rPheThrThrIleThrLeuProThrThrProThrSerPheThrThrThrThrThrThrTh 1287
 ||||| : : : : :
 QY 227 GACCCCAAGAGCCCTTCCGCCCGCCGCTTATCGAATCTGCGAACACAGCCTC 286
 ||||| : : : : :
 Db 1287 rThrProThr-----SerSerTh 1293
 ||||| : : : : :
 QY 287 GATTTGCAATACATATTGGCGCAACGACATTTCCGGCAGGCTCTGTTCACCTCA-- 344
 ||||| : : : : :
 Db 1293 rValLeuSerThrThr-----ProlLysLeuCysCysLeuTrpSerAs 1307
 ||||| : : : : :
 QY 344 ----- 344
 Db 1307 pTrpIleAsnGluAspHisProSerSerGlySerAspGlyAspArgGluProPheAs 1327
 ||||| : : : : :
 QY 345 -----TGTACCTTCCGCGAGACGCGCGGTGCGGTACGCC----- 380
 ||||| : : : : :
 Db 1327 pGlyValCysGlyAlaProGluAspIleGluCysArgSerValLysAspProHisLeuSe 1347
 ||||| : : : : :
 QY 380 ----- 380
 Db 1347 rLeuGluGlnHisGlyGlnLysValGlnCysAspValSerValGlyPheIleCysLysAs 1367
 ||||| : : : : :
 QY 380 ----- 380
 Db 1367 nGluAspGlnPheGlyAsnGlyProPheGlyLeuCysTyrAspTyrLysIleArgValAs 1387
 ||||| : : : : :
 QY 381 -TGCTGTTGG-----ACGACACACAC 400
 ||||| : : : : :
 Db 1387 nCysCysTrpProMetAspLysCysIleThrThrProSerProProThrThrThrProSe 1407
 ||||| : : : : :
 QY 401 ACGCGCGGTTGGACGATCTCTGCTCGCCCTCGACAGCCATCCCAATATCGAAGTGC 460

Pred. No.: 8,08e-16 Length: 5179
 Score: 336.50 Matches: 141
 Percent Similarity: 33.6% Conservative: 55
 Best Local Similarity: 24.1% Mismatches: 198
 Query Match: 11.8% Indels: 190
 DB: Gaps: 22

US-10-665-990a-13 (1-1561) x AEA27639 (1-5179)

QY	119	ACGGAAGCGTCATTCAAT	-----ACTTCAAAACCTGTCCTCGGACAC	166
DB	1248	ThrValGluLysHisPheAsnLeuCysSerIleThrThrArgPro-SerThrLeuThrTh	1267	
QY	167	ATCTCGAAATCGGCACACCCCTCATACACAGGCGTATCGACATCTACCTGCTCGAC	226	
DB	1267	rPheThrThrIleThrLeuProThrThrProThrSerPheThrThrThrThrTh	1287	
QY	227	GACCCACAGAGCCCTTCCCGCGCGCCCTTATCGAATCTGCGGAACACAGCCTC	286	
DB	1287	rThrProThr	1293	
QY	287	GATTTCGAATACACTACATTTGGCGCAACGACATTTCCGGCAGGCTGCTGTCAACCTCA	344	
DB	1293	rValLeuSerThrThr	1307	
QY	344	-----ProlLysLeuCysCysLeuTipSerAs	1307	
DB	1307	pTrrIleAsnGluAspHisProSerSerGlySerAspGlyAspArgGluProPheAs	1327	
QY	345	-----TGTACTTTCGCGCAGAACGCGCGGTGCGGTACGCC	380	
DB	1327	pGlyValCysGlyAlaProGluAspIleGluCysArgSerValLysAspProHisLeuSe	1347	
QY	380	-----ProlLysLeuCysCysLeuTipSerAs	1347	
DB	1347	rLeuGluGlnHisGlyGlnLysValGlnCysAspValSerValGlyPheIleCysLysAs	1367	
QY	380	-----ProlLysLeuCysCysLeuTipSerAs	1367	
DB	1367	nGluAspGlnPheGlyAsnGlyProPheGlyLeuCysTyrAspTyrLysIleArgValAs	1387	
QY	381	-TGCTCTTGG-----ACGACACAAAC	400	
DB	1387	nCysCysThrProMetAspLysCysIleThrThrProSerProProThrThrTh	1407	
QY	401	ACGCGCGGTGGACGATCTCTGCTCGCCTCGACAGCCATCCCAATATCGAAGTGC	460	
DB	1407	rProProProThrThrThrThrThrLeuProProThrThrPro	1422	
QY	461	CTGTTCAACCCCTTCTCTAGCAAAATGCGCGGCGCTCGGCTACCTGACCGACTCC	520	
DB	1423	-----SerProProThrThrThrThrThrThrThrThrThrThrThrTh	1441	
QY	521	CGCCTCAACCGCGCATGACACAAATCTTTACCGCGCAGCAACCGCGCCACCATCTC	580	
DB	1441	oleThrThr-----ThrThrProLeuProThrThrThrProSerPro-----	1456	
QY	581	GGCGGACGAATATCGCGACCAATATCTTCAAAGTCGGTGAGGACACCGTTTTCGCG	640	
DB	1457	-----ProlLysThrThrThrThrThrThrThrThrThrThrThrTh	1474	
QY	641	CTGGACATCTCGCCACCGCGAGCTGCTGCGGAGATATCGACGACTTCGAC	695	
DB	1474	rThrThrProProProProProProProProProProProProProProProTh	1494	
QY	696	-----GCTACTGGCAAGCATTCGCGCCACCAACCGCGCATCATCGCGAGCGC	748	
DB	1494	oProProThrThrThrProSerProProMetThrThrProlLysThrProProAlaSerTh	1514	
QY	749	AACATCGGAAGGCTTCAAGCACTCGGATACACGAAACATCCAGACAGCGGCTC	808	
DB	1514	rThrThrLeu-----ProProThrThrThrProSerProProThrThrTh	1529	

QY	809	CTGCGCTACCGGAAACCGTCGAAC-----AGTCGCCCTCTACCAAAATAACAGACG	862
DB	1529	rThrThrProProProThrThrThrProSerProProThrThrTh	1544
QY	863	GGACGATCGACTGGCAGAGCGTCAAAACCGCCTGATCAGCGACACCCCTGCAAAAGA	922
DB	1545	-----ThrProIleThrProPr	1550
QY	923	CTGACCGCGACCGCCGCAACCGCGGAGGCTGCAAGACGCGCTCAACAG	982
DB	1550	oThrSerThrThr-----LeuPro-----ProThrThrThrProSe	1563
QY	993	CCGAAAAAGCGCTATCTGTTTACCCCTATTTGCTCCCTACAAAATCGGCGACAG	1042
DB	1563	rProProProThrThrThrThrProProProThrThrThrProSerProProThrTh	1583
QY	1043	GCACTGGCAAAACTGGTCGACGCGCATAGACGTTACCGTCTCTGACCA-----AC	1093
DB	1583	r-----ThrThrProSerProProThrIleThrTh	1593
QY	1094	TGCTACAGGCGACGAGTTCGCCCGCTTCATTCGGCTACGTCAAT-----	1142
DB	1593	rThrThrProProProThrThrThrProSerProProThrThrThrThrProPr	1613
QY	1143	-----ACGAAAAACCGCTCAAAAGCGCGCATCAAACTC	1177
DB	1613	oProThrThrThrProSerProProThrThrThrProIleThrProThrSerThrTh	1633
QY	1178	TACGAGCTGCAACCCACCATGCTCCCGCCACAAAAGACAAAGGCTGACCGGACG	1237
DB	1633	rThrLeuProProProThrThrThrProSerProProProThrThr-----ThrThrTh	1650
QY	1238	TCCGTAACAGCGCTCATGCCAAAACCTTCATTTGGAGCGCAACGATCTTCATCGC	1297
DB	1650	rProProProThrThrThrProSerProProThr-----ThrThrThrProSerProPr	1669
QY	1298	TCATTCAACCTGACCCCGTTCGCGACGCTCAATACGAAATGGCGTCTCATCGAA	1357
DB	1669	eThrThrThrThrProProProThrThrThrThrProSerProIleThrThrThPr	1689
QY	1358	AGCCCCAAATGCGAGAACAGATGGAGCGCACCTTCGCCGATACACACCCGAAATCGC	1417
DB	1689	oSerProProThr-----ThrThrMetThrThrProSerProThrThrThrProSe	1708
QY	1418	TACCGCGTTCCTCGACAAACAGACCGCTGCAATGGCAGATCCCGCCACCCGAAA	1477
DB	1708	oile-----ThrThrThrThrProSerSerThrThrThrProSerProPr	1725
QY	1478	ACCTACCGAAGCAACCGCAACCTTTGGAACGATCGCGCAAAATCTATCC	1537
DB	1725	oThrThrMetThrThrProSerProThrThrThrProSerProThrThrMetTh	1745
QY	1538	CTGCTGCCCA 1547	
DB	1745	rThrLeuPro 1748	
RESULT 24			
ID	ABU53160		
XX	ABU53160	standard; protein; 395 AA.	
AC	ABU53160;		
XX	XX		
DT	14-APR-2003	(first entry)	
XX	XX		
DE	Human testes-derived DKFzphes3_2a11 homologue #20.		
XX	XX	Human; gene therapy; vaccine; disease treatment; detection.	
XX	XX	Homo sapiens.	
OS	XX		
PN	WO200112659-A2.		
XX	XX		
PD	22-FEB-2001.		

```
XX PF 18-AUG-2000; 2000WO-IB001496.
XX PR 18-AUG-1999; 99US-0149499P.
XX PR 28-SEP-1999; 99US-0156503P.
XX (GEHU-) GERMAN HUMAN GENOME PROJECT.
XX PA
XX WIemann S;
XX PI
XX WI
XX DR WPI; 2001-327840/34.
XX
XX Nucleic acids having the sequences of clones isolated from libraries of
XX PT different human tissues, useful in recombinant DNA methodologies.
XX PS
XX Example III; Page 774-775; 1095pp; English.
XX
XX This invention describes novel polynucleotides and polypeptides isolated
XX CC from human cDNA libraries which can be used for gene therapy or in
XX CC vaccines. The polynucleotides of the invention and antibodies encoded by
XX CC them may be used in the prevention, diagnosis and treatment of diseases
XX CC associated with inappropriate polypeptide expression. The products of the
XX CC invention may also be used to identify modulators of expression and
XX CC activity and to down regulate expression and activity. The antibodies of
XX CC the invention may also be used as diagnostic agents for detecting the
XX CC presence of polypeptides in samples. This sequence represents a homologue
XX CC of a polypeptide described in the disclosure of the invention
XX SQ Sequence 395 AA;

Alignment Scores:
Pred. No.: 1.83e-15 Length: 395
Score: 328.00 Matches: 148
Percent Similarity: 37.3% Conservative: 45
Best Local Similarity: 28.6% Mismatches: 179
Query Match: 11.5% Indels: 146
DB: 4 Gaps: 26

US-10-665-990A-13 (1-1561) x ABUS3160 (1-395)
QY 46 ACGCAGCTCATTTCCCTTTTATGCTCTCTCTGTTCTCATGTTCTTCATGGTGGCCCC 105
Db 2 ThrThrProSerProProThrThrThrProSerProProPro 15
QY 106 ACTGGAAGAACGGAGGAAAGCGTCATTTCAATACATTCCTCAAACTGTCCTCTCGGACAA 165
Db 16 ThrThrThrThr-----ThrLeuProProThrThr 25
QY 166 CATCTGCAATCCGGACACCCCTCATACACGGGCTATCGACATCTACCTGCTCG- 224
Db 26 ThrProSerProPro--ThrThrThrThrThrThrProProProProThrThrProSerP 45
QY 225 -----ACGACCCCGACGAGCCCTGGCG-----CCGCGCGCGCCCTTATCGA 267
Db 45 roProIleThrThrThrThrThrProLeuProThrThrThrProSerProProIleSerT 65
QY 268 ATCTCGCGAACACAGCCTCGATTTGCAATACTACATTTGGGGCAACGACATTTCCGGCAG 327
Db 65 hr-----ThrThrThrThrProProProProThrThrProPro-- 75
QY 328 GCTGCTGTTCACTGATGACCTGCGGACGAGACGGCGGTGGCGGTACGCGCTGCTGTT 387
Db 76 -----SerProProThrThrThrProSerProProThrThrThrProSerProProT 93
QY 388 GGACGACAAACACACGGCGGGTGGAGGATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 447
Db 93 hrThrThrThrThrThrProProProThrThrThrThrProSerProProMetThrThrProI 113
QY 448 TATCGAAGTGGCGCTGTTCAACC-----CCTTCGTCCTACGCAAAATGGCGGCGCACT 498
Db 113 leThrProProAlaSerThrThrThrLeuProProThrThr-----T 127
QY 499 CGGCTACCTGACCGACTTCCCGCGCCTCAACCGCGCGCATGCACAAACAAATCTTTACCGC 558
```

RESULT 25
AB073174


```
QY 481 ACGAAATGGCGGCACTCGGTACCTGACCGACTTCCCGCCCTCAACCGCGCATGCA 540
Db   ::::: |||||::: :::::
QY 124 -----ThrProSerProThrThrThrThrThrThrThrThrThrThrThrThr 138
QY 541 CAACAAATCCTTACCGCGCAACCC-----GCGCCACATACCTCGCGGCGCAA 591
Db   ||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 138 hrThrThrThrProSerProProThrThrThrProIleThrProProThrSerThrThrL 158
QY 592 TATCGCGCAGCAATCTTCAAGTCGGTGAGGACACCGTTTTCGCGACCTGGACATCCT 651
Db   :: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 158 euProProThrThrThrProSerPro-----ProProThrThrThrThrThr 173
QY 652 CGCACCGCGCAGCTCGTCGGGAGATATCGACGACTTCGACCTACTGGGCAAGCA 711
Db   ::||| ||| ::::: |||||::: |||||::: |||||::: |||||::: |||||:::
QY 173 hrProProProThrThrThrThrProSerProProThrThrThrThrProSerProThrI 193
QY 712 TTCGCGCCCAACAGCGCATCATCGCAGCGGCAACATCGCGGAGGCTCTTCAAGC 771
Db   || ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 193 leThrThrThrThrPro-----Pro----- 198
QY 772 ACTCGGATACAGCAAGCAATCCAGACACGCGTCTCGCTACCGCAAAACCGTGA 831
Db   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 199 --ProProThrThrThrProSerProProThrThrThrThrThrThrThrThrThr 218
QY 832 AC-----AGTCGCCCTCTACCAAAATATACAGAGCGGACCATCGACGAGCGT 885
Db   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 218 hrThrProSerProProThrThr----- 225
QY 886 CCAACCGCGCTGATCAGCGACCCCTGCAAAAGGACTCGACCGGACCGCGCAACCC 945
Db   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 226 -----ThrProIleThrProProThrThrThrThrThrThrThrThrThr 236
QY 946 GCCGATTCGCGGAGCTGCAAGACGCGCTCAACAGCCGCGCAAAAGCGTCTATCGT 1005
Db   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 237 -----LeuPro-----ProThrThrThrProSerPro----- 245
QY 1006 TTCACCTATTTCGTCCTTACAAATCCGGCACAGCGCACTGGCAAACTGGTGCAGGA 1065
Db   ||| ::::: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 246 -----ProProThrThrThrThrThrThrThrThrThrThrThrThrThrThr 256
QY 1066 CGGCATAGCTTACGCTCTGACCACTCGCTACAGGCGGACCGGCGCGTCA 1125
Db   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 257 -----ThrThrProSerProProThr-----ThrThrThrProSerP 269
QY 1126 TTCGCGTACGTCAAATACGCAAAACCGCTGCTCAAGCGCGCATCAAACTACGAGCT 1185
Db   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 269 roProIleThrThrThrThr-----ThrProProProThrThrThrProS 284
QY 1186 GCAACCCA-----ACCATGCGCTCCCGCCACAAAGACAAAGGCTGACCGGAGCTC 1239
Db   ::::: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 284 exSerProIleThrThrThrThrThrThrThrThrThrThrThrMetThrThrProSer 302
QY 1240 CGTAACACCGCTGATGCCAAAACCTTATTGTGGACGGCAACCGCATCTTCATCGGCTC 1299
Db   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 303 -----ProThrThrThrThrThrThrThrThrThrThrThrThrThrThr 318
QY 1300 ATTCAACTCGACCCCGTTCGCGACGCTCAATACCGAAATGGGCGTGCATCGAAG 1359
Db   ||||| ::::: ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 319 --SerThrThrThrProSerPro-----ProProThrThrMetThrThrProS 334
QY 1360 CCCCAAAATCGGAGACATGAGGCGCACCTTCGCGCATACACACCGCAATACGCTTA 1419
Db   ::||| ::: ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 334 exProThrThr-----ThrProSerProProThrThrThrThrMet 347
QY 1420 CGCGGTTACCTCGCAAAACACACCGCTGCAATGGCATCGGCGCACCGCAAAAC 1479
Db   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 347 hrThrLeuProProThrThrThrThrThrThrThrThrThrThrThrThrThrMet 367
QY 1480 CTACCGGAACGAAACCGAAGCAAACTTTGAAACGCATCGCGGCAAAATCTATTCCT 1539
Db   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 367 leThrProProThrPheSerPro---PheSerThrThrThrThrPro-----ThrThrProC 384
QY 1540 GCTGCCCA 1547
```

```
Db 384 ysValPro 386
RESULT 27
ABU53144
ID ABU53144 standard; protein; 717 AA.
XX
AC ABU53144;
XX
DT 14-APR-2003 (first entry)
XX
DE Human testes-derived DKFZphtes3_2all1 homologue #4.
XX
KW Human; gene therapy; vaccine; disease treatment; detection.
XX
OS Homo sapiens.
XX
PN WO200112659-A2.
XX
PD 22-FEB-2001.
XX
PF 18-AUG-2000; 2000WO-IB001496.
XX
PR 18-AUG-1999; 99US-0149499P.
PR 28-SEP-1999; 99US-0156503P.
XX
PA (GEHU-) GERMAN HUMAN GENOME PROJECT.
XX
PI Wiemann S;
XX
WPI; 2001-327840/34.
XX
PT Nucleic acids having the sequences of clones isolated from libraries of
different human tissues, useful in recombinant DNA methodologies.
XX
PS Example III; Page 765; 1095pp; English.
XX
CC This invention describes novel polynucleotides and polypeptides isolated
from human cDNA libraries which can be used for gene therapy or in
vaccines. The polynucleotides of the invention and antibodies encoded by
them may be used in the prevention, diagnosis and treatment of diseases
associated with inappropriate polypeptide expression. The products of the
invention may also be used to identify modulators of expression and
activity and to down regulate expression and activity. The antibodies of
the invention may also be used as diagnostic agents for detecting the
presence of polypeptides in samples. This sequence represents a homologue
of a polypeptide described in the disclosure of the invention
XX
SQ Sequence 717 AA;
Alignment Scores:
Pred. No.: 3,52e-15 Length: 717
Score: 325.00 Matches: 140
Percent Similarity: 38.2% Conservative: 45
Best Local Similarity: 28.9% Mismatches: 169
Query Match: 11.4% Indels: 130
DB: 4 Gaps: 24
US-10-665-990A-13 (1-1561) x ABU53144 (1-717)
QY 148 ACCTGCTCTCTCGGCAACATCTCGAAATCGGCACACCCCTCATACACAGCGGCTATC 207
Db 15 ThrLeuProProThrThrThrThrProSerProPro--ThrThrThrThrThrThrPro 34
QY 208 CGACATCTACCTGCTCG-----ACGACCCCAACGAGCCCTTGCGG-----C 249
Db 34 roThrThrThrProSerProIleThrThrThrThrThrProLeuProThrThrThrP 54
QY 250 CGCGCGCCCTTATCGAATCTGCGCAACACAGACCTCGATTTGCAATACATATTGGCG 309
Db 54 roSerProProIleSerThr-----ThrThrThrThrThrThrThrThrThrThr 66
QY 310 CAACGACATTTCCGGCAGGCTGCTGTCAACCTCATGTACCTTGGCGCAAGACGCGCGT 369
```


Qy	904	CGACACCCCTGC	AAAGGACTCG	ACGCGCGG	CGA	CCGCGCAAA	CCGCGCAATTC	CGCGGAGGCT	963
Db	244	--ThrProIle	ThrProThr	SerThrThr	ThrThr	-----	LeuPro	-----	256
Qy	964	GCAGACGCGCT	CAACACAGC	CGCCCAAAA	AGGCTG	TATCTG	TTTTCACCT	ATTTCGTCC	1023
Db	257	--ProThrThr	ThrProSer	Pro-----	-----	-----	ProProThr	ThrT	268
Qy	1024	TACAAAAT	CGGCACAG	ACGACCTG	GCAAACTG	GTG	CAGGACG	CGCATACGCTT	1083
Db	268	hrThrThr	ProProPro	Thr-----	-----	-----	ThrThrPro	S	278
Qy	1084	CTTGACCACT	CGCTACAG	CGACGCG	AGCTTG	CGCGCG	CTCCATTC	CGCGTAGCT	1143
Db	278	erProProThr	-----	-----	ThrThrThr	ThrThrThr	ProIleThr	ThrThrT	293
Qy	1144	CCGAAAAC	CGCTGCT	CAAGCCGG	ATCAAACT	CTC	TAGGAGCT	GCAACCCA	1197
Db	293	hr-----	-----	-----	ThrProPro	ThrThrThr	ProSerSer	ProIleThrThrT	308
Qy	1198	TGCGGTCC	CGGCACAAA	AGAGCGC	CTGAC	CGGCAGCT	CCGTAA	CCAGCGCTG	1257
Db	308	hrProSer	ProProThr	ThrThrMet	ThrThrPro	Ser-----	-----	ProThrThrThrP	325
Qy	1258	CAAAAC	CTTCAT	TGTGGACG	CGAAACGC	ATCTTC	ATFCGGCT	CATTCAAC	1317
Db	325	roSerSer	ProIleThrThrThrThrThr	ThrThrPro	Ser-----	-----	SerThrThrThr	ProS	342
Qy	1318	TTCCGACG	CGCTCA	ATACGAA	TGGCGT	CGTCAT	CGAAAGC	CCCAAAAT	1377
Db	342	erPro-----	-----	-----	ProProThr	ThrMetThrThr	ProSerPro	ThrThrT	355
Qy	1378	GATGGAGG	CACCTCG	CGATACCA	CACCGGA	TACGCTAC	CGCGTTAC	CCCTCGACA	1437
Db	356	-----	-----	-----	ThrProSer	ProProThrThrThrMet	ThrThrLeu	ProProThrT	371
Qy	1438	ACACAA	CCGCGCTG	CAATGC	AGATCC	CGCCAC	CCGAAAAA	CACTACC	1484
Db	371	hrThrSer	SerProLeu	ThrThrThrPro	LeuProPro	SerIleThr	386		
RESULT	30								
ABU53157									
ID	ABU53157	standard; protein; 368 AA.							
XX	ABU53157;								
XX	AC								
XX	14-APR-2003	(first entry)							
XX	Human testes-derived DKF2phtes3_2a11	homologue #17.							
XX	Human; gene therapy; vaccine; disease treatment; detection.								
XX	Homo sapiens.								
XX	WO200112659-A2.								
XX	22-FEB-2001.								
XX	18-AUG-2000; 2000WO-IB001496.								
XX	18-AUG-1999; 99US-0149499P.								
XX	28-SEP-1999; 99US-0156503P.								
XX	(GEHU-) GERMAN HUMAN GENOME PROJECT.								
PA	Wiemann S;								
PI	WPI; 2001-327840/34.								
XX	XX								
XX	XX								
XX	XX								
PT	Nucleic acids having the sequences of clones isolated from libraries of								
PT	different human tissues, useful in recombinant DNA methodologies.								
XX	Example III; Page 773-774; 1095pp; English.								
PS	XX								

Db	219	hrThrProSerProProThrThr-----	226
Qy	886	CCAAACCCGCTGATCAGCGACACCCCTGAAAGAACTCGACGGACGCGCGCAAAACC	945
Db	227	-----ThrProIleThrProProThrSerThrThr-----	237
Qy	946	GCCGATTGCGGAGGCTGCAAGACCGGCTCAAAACAGCCCGGAAAGAGCTCTATCTGGT	1005
Db	238	-----LeuPro-----ProThrThrThrProSerPro-----	246
Qy	1006	TTCACCCCTATTTGCTCCCTACAAAATCCGCACAGACGCTGGCAAAACTGGTCAGGA	1065
Db	247	-----ProProThrThrThrThrProProProThr-----	257
Qy	1066	CGGCATAGAGCTTACCGTCTGACCAACTCGCTACAGGCGACCGAGTGGCCGCGCTCCA	1125
Db	258	-----ThrThrProSerProProThr-----ThrThrThrProSerP	270
Qy	1126	TTCCGGTACGCTCAAAATACCGAAACCGGTGCTCAAAGCGCGCATCAAACTTACAGCT	1185
Db	270	roProlleThrThrThrThr-----ThrProProProThrThrThrPros	285
Qy	1186	GCAACCCA-----ACATGCGCTCCCGGCCACAAAGACAAAGGCTGACCGCAGCTC	1239
Db	285	erSerProlleThrThrThrProSerProProThrThrThrMetThrThrProSer----	303
Qy	1240	CGTAACCAAGCTGCTGTCGCCAAAACCTTCATTGTGGACGGCAACGCGATCTTCATCGGCTC	1299
Db	304	-----ProThrThrThrProSerSerProIleThrThrThrThrProSer-----	319
Qy	1300	ATTCAACCTCGAACCCCGTTTCGGACGGCTCAATACGAAATGGGCGTCTCATCGAAG	1359
Db	320	--SerThrThrThrProSerPro-----ProProThrThrMetThrThrPros	335
Qy	1360	CCCCAAAATCGCAGAACAGATGGAGCGCACCTCGCGATACCAACCCCGAAATACGCCTA	1419
Db	335	erProThrThr-----ThrProSerProProProThrThrThrMetT	348
Qy	1420	CCGCGTTACCTCGACAAACAAACCGCTGCAATGGCAGCATCCGCGCAACC	1472
Db	348	hrThrLeuProProThrThrThrSerSerProLeuThrThrThrProLeuPro	365

RESULT 31	
ABM71816	
ID	ABM71816 standard; protein; 513 AA.
XX	
AC	ABM71816;
XX	
DT	20-NOV-2003 (first entry)
XX	
DE	Staphylococcus aureus protein #1056.
XX	
KW	Antibacterial; vaccine; gene therapy; infection; sepsis; diagnosis;
KW	enzymatic assay; antibiotic target.
XX	
OS	Staphylococcus aureus.
XX	
PN	WO200294868-A2.
XX	
PD	28-NOV-2002.
XX	
PF	27-MAR-2002; 2002WO-IB002637.
XX	
PR	27-MAR-2001; 2001GB-00007661.
XX	
PA	(CHIR-) CHIRON SPA.
XX	
PI	Masignani V, Mora M, Scarselli M;
XX	
DR	WPI; 2003-120786/11.
DR	N-PSDB; ACF73376.
XX	

Db	373	SerProGluLysLeuProGlnSerSerSerSerGluSerSerProProSerPro	---Gln	391
Qy	250	CCGCGCGCCCTTATCGAATCTCCGAAACACACGCTCGATTTCGAACTACTACATTTGGCG	309	
Db	392	ProThrLysValSerArgHisAlaSerSerSerProGluSerProLysProAlaProAla	411	
Qy	310	CAACGACATTTCCGGCAGCGCTGCTTTCAACCT	342	
Db	412	ProGlySerHisArgGluLeuSerSerSerProThrSerLysAsnArgSerHisGlyArg	431	
Qy	343	-----CATGTACCTTCGCGAGAACGCGCGCTCGCGCTGACGTACCT	378	
Db	432	AlaLysArgAspLysSerHisThrProSerArgArgMetGlyArgSerArgSer	451	
Qy	379	CTGCTGTTGGACGACAACAACCGCGGGTTGGACGATCTCTCT	423	
Db	452	ProAlaThrAlaLysArgGlyArgSerArg	470	
Qy	424	GCTCGCCCTCGACAGCATCCCAATATCGAAGTGCCTGTTCAACCCCTTCGTCCTTACG	483	
Db	471	SerArgSerArgSerProGlnTrpArgSerArgSerAlaGln	485	
Qy	484	CAAAATCGCGCGACTCGGCTACTGACCGACTTCCCGCGCTCAACCGCGCATGCACAA	543	
Db	486	ArgTrpGlyArgSerArgSerProGlnArgArgGlyArgSerPro	502	
Qy	544	CAAAATCCTTTACCGCGCAGACCGCGCCACCATTCTCGGGGAGCGCAATATCGCGACGA	603	
Db	503	-----GlnArgProGlyTrpSerArgSerArgAsnThrGlnArgArgGlyArg	518	
Qy	604	ATACTTCAAAGTCGGTGAGGACACCGTTTTCGCGGACCTGGACATCTCGCCACCCGCGCAG	663	
Db	519	SerArgSerAlaArgGlyArgSerHisSerArgSerProAlaThrArgGlyArgSer	538	
Qy	664	CGTGTGCGGGAAGTATCGACGACTTCGACCGCTACTGGGCAAGCATTCGCGCCACAA	723	
Db	539	ArgSerArgThrProAlaArgArgGlyArgSerArgSerArgThrProAlaArgArgArg	558	
Qy	724	CGCACGCGCATCATCGGAGCGGCAACAT	762	
Db	559	SerArgSerThrProThrArgArgArgSerArgSerArgThrProAlaArgArgGly	578	
Qy	763	TCTTCAAGCATCGGATA---CAAACGACGAAACATCCAGACACGCGCTCTCGCTACCG	819	
Db	579	ArgSerArgSerArgThrProAlaArgArgSerArgSerArgThrArgSerPro---ValArg	597	
Qy	820	CGAAACCGTTCGACAGTGCCTCTACCAAAANAATACAGACGGGCGCATCGCTGCCA	879	
Db	598	ArgArgSerArgSerArgSerProAlaArgArgSerGlyArgSerArgSerArgThrPro	617	
Qy	880	GAGCGTCCAAACCGCTGATCAGCGACACCCCTGCAAA-----AGGACTCGACCGCGA	933	
Db	618	AlaArgArgGlyArgSerArgSerArgThrProAlaArgArgGlyArgSerArgSerArg	637	
Qy	934	CCGCGCGAAACCCCGATTCGCGGAGGCTGCAAGA-----	969	
Db	638	ThrProAlaArgArgSerGlyArgSerArgSerArgThrProAlaArgArgGlyArgSer	657	
Qy	970	CGCGCTCAAAACCGCGAAGAAACGCTATCTGTTTTCACCCCTATTTCCTCCCTACAAA	1029	
Db	658	ArgSerArgThrProArgArgGlyArgSerArgSerArgSerLeuValArgArgGlyArg	677	
Qy	1030	ATCCGGCACAGA-----CGCACTGGCAAAATGTTGCGAGCAGCGCATAGACGTTAC	1080	
Db	678	SerHisSerArgThrProGlnArgArgGlyArgSerGlySerSerSerGluArgLys---	696	
Qy	1081	CGTCTGACCACTCGCTACAGCGGACCGACGTTGCGC-----	1119	
Db	697	-----AsnLysSerArgThrSerGlnArgArgSerArgSerAsnSerSerProGluMet	714	
Qy	1120	-----CGTCCATTTCGCGCTACGTCAAAATACCGAAACCGCTGCT	1158	

Alignment Scores:					
Pred. No.:	1.45e-13	Length:	956		
Score:	303.50	Matches:	156		
Percent Similarity:	37.6%	Conservative:	78		
Best Local Similarity:	25.1%	Mismatches:	233		
Query Match:	10.6%	Indels:	156		
DB:	6	Gaps:	29		
US-10-665-990A-13 (1-1561) x ABO53053 (1-956)					
Qy	43	AACAGCGAGCTCATTTCCCTTTTANGCCTCTCTCTGTCATGTTCTTCANGTTGCC	102		
Db	323	SerThrGlyProGluProProAlaProThrProLeuLeu-	335		
Qy	103	CCCAGTGAAAGAACGGACGGAAAGCCGTCAATTCAATACTTCCAACCTGTCTCTCTGGA	162		
Db	336	--AlaGluArgHisGlySerPro-----GlnProLeuAlaThrThrProLeuSer	352		
Qy	163	CAACATCTCGCAATTCG-----GCACACCCTCATACAACAGGCTATCGACAT	213		
Db	353	GlnGluProValAsnProProSerGluAlaSerProThrArgAspArgSerProProlys	372		
Ov	214	CTACTGTGTCGACGACCC-----CCACGAAGCCCTCTGCCGC	249		

Db 715 LysLysSerArgIleSerSerArgSerArgSerLeuSerSerProArgSerLysala 734
 QY 1159 CAAGCCGGCATCAAACTCTACGA-----GCTGCAACCCCAACATGCGTCCCGCCAC 1212
 Db 735 LysSerArgLeuSerLeuArgSerLeuSerGlySerSerProCys---ProLysGln 753
 QY 1213 AAAGACAAAGCCGTGACGGCAGCTCCGT-----AACAGCCTGC----- 1253
 Db 754 LysSerGlnThrProProArgSerArgSerGlySerGlnProLysAlaLysSer 773
 QY 1254 -----ATGCCAAACCTTCATTGTTGGACGCCA 1280
 Db 774 ArgThrProProArgSerArgSerSerSerProProLysGlnLysSerLys 793
 QY 1281 AACGCATCTTCATCGGCTCATTCAACCTCGACCCCGCTTCG----- 1322
 Db 794 ThrProSerArgGlnSerHisSerSerSerProHisProLysValLysSerGlyThr 813
 QY 1323 -----CAGGCTCAATACCGAATGGCGTGCATCGAAGCCCAAAATCGCAG 1373
 Db 814 ProProArgGlnGlySerIleThrSerProGlnAlaAenGluGlnSerValThrProGln 833
 QY 1374 AACAGATGAGCGACCCCTCCGCGATACACACCGAATACGCTACCGCGTTACCTCG 1433
 Db 834 ArgArg---SerCysPheGluSerSerProAspProGluLeuLysSerArgThrProSer 852
 QY 1434 ACAACACAAACCGCTCGAATGGC-----ACGATCCCGCCA 1469
 Db 853 ArgHis-----SerCysSerGlySerSerProProArgValLysSerSerThrProPro 870
 QY 1470 CCGAATAACTACCGCA-----ACGAACCCGAAGCCA----- 1502
 Db 871 ArgGlnSerProSerArgSerSerSerProGlnProLysValLysAlaIleSerPro 890
 QY 1503 -----AATTGGAAACGATCGCGCGCAAAATCCTATCCTGCTGCCATCG 1550
 Db 891 ArgGlnArgSerHisSerGlySerSerSerProSer-----ProSer 904
 QY 1551 AAGGTT 1556
 Db 905 ArgVal 906
 RESULT 33
 ADN60280
 ID ADN60280 standard; protein; 2296 AA.
 XX
 AC ADN60280;
 XX
 DT 18-NOV-2004 (first entry)
 XX
 DE Human splicing coactivator subunit SRM300 protein.
 XX
 KW angiogenesis modulator; tumorigenesis modulator; angiogenesis; Ax1;
 KW tubulin cofactor D; transglutaminase 2; cytosine deaminase;
 KW peptidase M41; paraplegin; CD13 aminopeptidase; PPK-1; zip kinase; Gas6;
 KW SRM160; non-muscle myosin heavy chain; calmodulin 2; symporter;
 KW semaphorin; zinc finger helicase; plexin-A2; deoxycytidylate deaminase;
 KW sugar transporter; tumorigenesis; antiangiogenic; cytostatic;
 KW cerebroprotective; vasotropic; antiinfertility; cardiant;
 KW antibody therapy; antisense therapy; RNA interference therapy;
 KW RNAi therapy; cancer; stroke; infertility; heart disease; human.
 XX
 OS Homo sapiens.
 XX
 XX W02004039955-A2.
 PN 13-MAY-2004.
 XX
 PD 29-OCT-2003; 2003WO-US034281.
 XX
 PF 29-OCT-2002; 2002US-0421989P.
 XX
 PR 17-OCT-2003; 2003US-0512251P.
 XX

(RIGE-) RIGEL PHARM INC.

Lorens JB, Atchison RE, Frieria A, Holland S;

WPI; 2004-376181/35.

N-ESDB; ADN60279.

Identifying a compound that modulates angiogenesis or tumorigenesis, useful in diagnosing and treating angiogenesis, cancer, stroke, infertility and heart disease, comprises contacting the compound with angiogenesis polypeptide.

Disclosure; Page 70-71; 105pp; English.

The present invention describes a method for identifying a compound that modulates angiogenesis or tumorigenesis. The method comprises: (a) contacting the compound with angiogenesis polypeptide, e.g. Ax1, tubulin cofactor D, transglutaminase 2, cytosine deaminase, peptidase M41 (paraplegin), CD13 aminopeptidase, PPK-1, zip kinase, Gas6, SRM160, non-muscle myosin heavy chain, calmodulin 2, novel symporter, novel semaphorin, novel zinc finger helicase (FLJ22611), plexin-A2, deoxycytidylate deaminase or novel sugar transporter; (b) determining the functional effector of the compound upon the angiogenesis polypeptide or the physical effect of the compound upon the target polypeptide or its fragment or inactive variant; and (c) determining the chemical or phenotypic effect of the compound upon a cell comprising the target polypeptide or its fragment or inactive variant, thus identifying a compound that modulates cell cycle arrest. Also described is a method of modulating angiogenesis in a subject. The angiogenesis or tumorigenesis modulating compound has antiangiogenic, cytostatic, cerebroprotective, vasotropic, antiinfertility and cardiant activities, and can be used in antibody, antisense and RNA interference (RNAi) therapies. The method is useful in identifying a compound that modulates angiogenesis. The methods and compounds or compositions are useful in diagnosing and treating angiogenesis, cancer, stroke, infertility and heart disease. The present sequence represents a human splicing coactivator subunit SRM300 protein, CC which is used in the exemplification of the present invention.

XX Sequence 2296 AA;

Alignment Scores:

Pred. No.:	1,79e-13	Length:	2296
Score:	303.50	Matches:	156
Percent Similarity:	37.6%	Conservative:	78
Best Local Similarity:	25.1%	Mismatches:	233
Query Match:	10.6%	Indels:	156
DB:	8	Gaps:	29

US-10-665-990A-13 (1-1561) x ADN60280 (1-2296)

QY 43 AACACGAGCCTCATTTCCCTTTTATGCTCTCTCTCATGTCTTCATGGTTGCC 102
 Db 358 SerThrGlyProGluProProAlaProThrProLeuLeu----- 370
 QY 103 CCCACTGGAAGAACGACGAGGAAAGCGCTCATTTCAATACTTCCAAACCTGCTCCTCGGA 162
 Db 371 ---AlaGluArgHisGlyGlySerPro-----GlnProLeuAlaThrProLeuSer 387
 QY 163 CAACATCTCTGCAAAATCCG-----GCACACCTCTATAACACGGGTATCCGACAT 213
 Db 388 GlnGluProValAsnProProSerGluAlaSerProThrArgAspArgSerProProLys 407
 QY 214 CTACCTGCTCGACGACCC-----CCACGAACGCTTGGCGC 249
 Db 408 SerProGluLysLeuProGlnSerSerSerGluSerSerProProSerPro---Gln 426
 QY 250 CCGCGCGCCCTTATCGAATCTCCGGAACACACCTCGATTTCATACATACATTGGCG 309
 Db 427 ProThrLysValSerArgHisAlaSerSerSerProGluSerProAlaProAla 446
 QY 310 CAACGACATTTCCGCGAGGCTGCTGTTCAACCT----- 342
 Db 447 ProGlySerHisArgGluIleSerSerSerProThrSerLysAsnArgSerHisGlyArg 466

QY 343 -----CATGTACCTTCCCGCAGAACCGCGGTCCGGTACG 378
 Db 467 AlalysArgAspLysSerHisSerHisThrProSerArgArgMetGlyArgSerArgSer 486
 QY 379 CCTGCTGTGGACGACAAACACACCGCGGGTGGACGATCCCT- 423
 Db 487 ProAlaThrAlaLysArgGlyArgSerArg---SerArgThrProThrLysArgGlyHis 505
 QY 424 GTCGCTCTCGACAGCCATCCCAATATCGAAGTCGCCTGTGTCAACCCCTTCGTCTACG 483
 Db 506 SerArgSerArgSerProGlnTrpArgArgSerArgSerAlaGln----- 520
 QY 484 CAAATGGCGCGACTCGGTACTACCGACTTCCCGCCCTCAACCGCGCATCCACAA 543
 Db 521 ArgTrpGlyArgSerArgSerProGlnArgGlyArgSerArgSerPro----- 537
 QY 544 CAAATCCTTTACCGCGCAACCGCGCCACCATCTCGCGCGACCAATATCGGCGACGA 603
 Db 538 -----GlnArgProGlyTrpSerArgSerArgAsnThrGlnArgArgGlyArg 553
 QY 604 ATACTTCAAAGTCGTGAGCACCCGTTTTCGCCGACCTGGACATCTCTCGCCACCGGAG 663
 Db 554 SerArgSerAlaArgGlyArgSerHisSerArgSerProAlaThrArgGlyArgSer 573
 QY 664 CGTCGTCGGGAGATATCGCAGACTTCGACCGCTACTGGGCAAGCCATTCCGCCACAA 723
 Db 574 ArgSerArgThrProAlaArgArgGlyArgSerArgSerArgThrProAlaArgArg 593
 QY 724 CGCCACGGCATCATCCGCGCGCAACAT-----CGGCAAGGG 762
 Db 594 SerArgSerArgThrProThrArgArgSerArgSerArgThrProAlaArgArgGly 613
 QY 763 TCTTCAAGCACTCGGATA---CAACGACGAAACATCCAGACACGCGCTCTCGGTACCG 819
 Db 614 ArgSerArgSerArgThrProAlaArgArgSerArgSerArgThrArgSerPro---ValArg 632
 QY 820 CGAACCCTCGAACAGTCGCCCTCTACCAAAAATACAGACGGACGCATCGATGGCA 879
 Db 633 ArgArgSerArgSerArgSerProAlaArgSerGlyArgSerArgSerArgThrPro 652
 QY 880 GAGCGTCCAAACCGCCTGATCAGCGACACCCCTGCAAA-----AGGACTCGACCGCA 933
 Db 653 AlaArgArgGlyArgSerArgSerArgThrProAlaArgArgGlyArgSerArgSerArg 672
 QY 934 CGCGCGCAACCGCGCATTCGCGAGGCTGCAAGA----- 969
 Db 673 ThrProAlaArgArgSerGlyArgSerArgSerArgThrProAlaArgArgGlyArgSer 692
 QY 970 CGCGCTCAACACCGCGAAAAGCGTCTATCTGGTTTCACCTATTTCGTCCCTACAA 1029
 Db 693 ArgSerArgThrProArgArgGlyArgSerArgSerArgSerLeuValArgGlyArg 712
 QY 1030 ATCCGCGCACAGA-----CGCACTGGCAAACTGGTGCAGACGGCATAGACGTTAC 1080
 Db 713 SerHisSerArgThrProGlnArgArgGlyArgSerGlySerSerSerGluArgLys--- 731
 QY 1081 COTCTGACCAACTCGCTACAGCGACCGACGCGTTCGCGC----- 1119
 Db 732 -----AsnLysSerArgThrSerGlnArgArgSerArgSerAsnSerSerProGluMet 749
 QY 1120 -----CGTCCATTCCGCTACGTCAAATACCGAAAACCGCTGCT 1158
 Db 750 LysLysSerArgIleSerArgSerArgSerLeuSerArgSerProArgSerLysAla 769
 QY 1159 CAAAGCGCGCATCAAACTCTACGA-----GTCGAACCCACCATCGCTCCCGCCAC 1212
 Db 770 LysSerArgLeuSerLeuArgArgSerLeuSerGlySerProCys---ProLysGln 788
 QY 1213 AAAAGACAAAGCGCTGACCGGAGCTCCGT-----AACGAGCTGC----- 1253
 Db 789 LysSerGlnThrProProArgArgSerArgSerGlySerGlnProLysAlaLysSer 808

QY 1254 -----ATGCCAAACCTTCATTGTGGACGGCA 1280
 Db 809 ArgThrProProArgArgSerArgSerSerSerProProLysGlnLysSerLys 828
 QY 1281 AACGCATCTTCATCGGCTCATTCACCTCGACCCCGTTCGC- 1322
 Db 829 ThrProSerArgGlnSerHisSerSerSerProHisProLysValLysSerGlyThr 848
 QY 1323 -----CAGCGCTCAATACCGAATGGCGGTGCTCATCGAAAGCCCCAATCGCAG 1373
 Db 849 ProProArgGlnGlySerIleThrSerProGlnAlaAsnGlnSerValThrProGln 868
 QY 1374 AACAGATGAGCGCACCCCTCCCGCATACACACCGCAATACCGCTACCGCGTTACCCCTCG 1433
 Db 869 ArgArg---SerCysPheGluSerSerProAspProGluLeuLysSerArgThrProSer 887
 QY 1434 ACAAAACAACCGCTCGCAATGGC-----ACGATCCCGGCA 1469
 Db 888 ArgHis-----SerCysSerGlySerSerProProArgValLysSerSerThrProPro 905
 QY 1470 CCGGAAAACCTACCCGA---ACGAACCCGGAAGCA----- 1502
 Db 906 ArgGlnSerProSerArgSerSerProGlnProLysValLysAlaIleSerPro 925
 QY 1503 -----AACTTTGAAACGATCGCGCGCAAAAATCCTATCCCTGCTGCCCATCG 1550
 Db 926 ArgGlnArgSerHisSerGlySerSerProSer-----ProSer 939
 QY 1551 AAGGTT 1556
 Db 940 ArgVal 941
 RESULT 34
 ID ADN60278 standard; protein; 2752 AA.
 XX AC ADN60278;
 XX DT 18-NOV-2004 (first entry)
 XX DE Human serine/arginine repetitive matrix 2 protein.
 KW angiogenesis modulator; tumorigenesis modulator; angiogenesis; Axl;
 KW tubulin cofactor D; transglutaminase 2; cytosine deaminase;
 KW peptidase M41; paraplegin; CD13 aminopeptidase; PK-1; zip kinase; Gas6;
 KW SRM160; non-muscle myosin heavy chain; calmodulin 2; symporter;
 KW semaphorin; zinc finger helixase; plexin-A2; deoxycytidylate deaminase;
 KW sugar transporter; tumorigenesis; antiangiogenic; cytosolic;
 KW cerebroprotective; vasotropic; antifertility; cardiac;
 KW antibody therapy; antisense therapy; RNA interference therapy;
 KW RNAi therapy; cancer; stroke; infertility; heart disease; human.
 OS Homo sapiens.
 XX WO2004039955-A2.
 XX PD 13-MAY-2004.
 XX PF 29-OCT-2003; 2003WO-US034281.
 XX PR 29-OCT-2002; 2002US-0421989P.
 XX PR 17-OCT-2003; 2003US-0512251P.
 XX (RIGE-) RIGEL PHARM INC.
 XX Lorenz JB, Atchison RE, Frier A, Holland S;
 XX WPI: 2004-376181/35.
 XX DR N-PSDB; ADN60277.
 PT Identifying a compound that modulates angiogenesis or tumorigenesis,
 PT useful in diagnosing and treating angiogenesis, cancer, stroke,
 PT infertility and heart disease, comprises contacting the compound with

angiogenesis polypeptide.

Disclosure; Page 68; 105pp; English.

The present invention describes a method for identifying a compound that modulates angiogenesis or tumorigenesis. The method comprises: (a) contacting the compound with angiogenesis polypeptide, e.g. Axl, tubulin cofactor D, transglutaminase 2, cytosine deaminase, peptidase M41 (paraplegin), CD13 aminopeptidase, PPK-1, zip kinase, Gas6, SRm160, non-muscle myosin heavy chain, calmodulin 2, novel symporter, novel semaphorin, novel zinc finger helicase (FLJ22611), plexin-A2, deoxycytidylate deaminase or novel sugar transporter; (b) determining the functional effector of the compound upon the angiogenesis polypeptide or the physical effect of the compound upon the target polypeptide or its fragment or inactive variant; and (c) determining the chemical or phenotypic effect of the compound upon a cell comprising the target polypeptide or its fragment or inactive variant, thus identifying a compound that modulates cell cycle arrest. Also described is a method of modulating angiogenesis in a subject. The angiogenesis or tumorigenesis modulating compound has antiangiogenic, cytostatic, cerebroprotective, vasotropic, antiinfertility and cardiact activities, and can be used in antibody, antisense and RNA interference (RNAi) therapies. The method is useful in identifying a compound that modulates angiogenesis. The methods and compounds or compositions are useful in diagnosing and treating angiogenesis, cancer, stroke, infertility and heart disease. The present sequence represents a human serine/arginine repetitive matrix 2 protein, which is used in the exemplification of the present invention.

Sequence 2752 AA;

Alignment Scores:

Pred. No.: 1.87e-13 Length: 2752
Score: 303.50 Matches: 156
Percent Similarity: 37.6% Conservative: 78
Best Local Similarity: 25.1% Mismatches: 233
Query Match: 10.6% Indels: 156
DB: 8 Gaps: 29

US-10-665-990A-13 (1-1561) x ADN60278 (1-2752)

QY 43 AACACGAGCTCATTTCCCTTTATGCTCTCTCTGTCATGTTCTTCATGGTGCC 102
DB 358 SerThrGlyProGluProAlaProThrProLeuLeu----- 370
QY 103 CCCACTGGAAGACGCGAAGCGCTCATTTCAATACTTCCAAACCTGCTCCTCTGGA 162
DB 371 ---AlaGluA9HisGlySerPro-----GlnProLeuAlaThrProLeuSer 387
QY 163 CAACATCTCGAAATCCG-----GCACACCCCTCATAAACGGGCTATCCGACAT 213
DB 388 GlnGluProValAsnProProSerGluAlaSerProThrArgAspArgSerProProlys 407
QY 214 CTACTGCTCAGACACC-----CCACGAGCCCTTGCCGC 249
DB 408 SerProGluLysLeuProGlnSerSerSerSerGluSerSerProProSerPro---Gln 426
QY 250 CCGCGCCCTTATCGAATCTCGGACACAGCTCGATTGCAATACATTTGGCG 309
DB 427 ProThrLysValSerArgHisAlaSerSerSerSerProGluSerProLysProAlaProAla 446
QY 310 CAACGACATTTCCGGCAGGCTGCTGTCAACCT----- 342
DB 447 ProGlySerHisArgGluLeuSerSerSerProThrSerLysAsnArgSerHisGlyArg 466
QY 343 -----CATGTACTTCCGCGACAGACGGCGGTGCGGTACG 378
DB 467 AlaLysArgAspLysSerHisSerHisThrProSerArgArgMetGlyArgSerArgSer 486
QY 379 CCTGCTGTGGACGACACACACGCGGGTGGACGATCTCCT----- 423
DB 487 ProAlaThrAlaLysArgGlyArgSerArg---SerArgThrProThrLysArgGlyHis 505
QY 424 GCTCGCCCTCGACGCCATCCCAATATCGAAGTGGCGCTGTTCACCCCTTCGTCCTACG 483

DB 506 SerArgSerArgSerProGlnTrpArgArgSerArgSerAlaGln----- 520
QY 484 CAATATGGCGCCACTCGCTACTGACCGACTTCCCGCTCAACCCGCGCATGCACAA 543
DB 521 ArgTrpGlyArgSerArgSerProGlnArgArgGlyArgSerArgSerPro----- 537
QY 544 CAATATCTTTACCCGCGACAAACCGCCACCATATCTCGGCGAGCAATATCGCGACGA 603
DB 538 -----GlnArgProGlyTrpSerArgSerArgAsnThrGlnArgGlyArg 553
QY 604 ATACTTCAAGTCGGTGAGGACACCGTTTTCGCGACCTGGACATCTCTCGCCACCGGAG 663
DB 554 SerArgSerAlaArgGlyArgSerHisSerArgSerProAlaThrArgGlyArgSer 573
QY 664 CGTCGTGGCGAAGTATCGCAGCACTTCGACCGCTACTGGCAAGCCATTCGCGCCACAA 723
DB 574 ArgSerArgThrProAlaArgGlyArgSerArgSerArgThrProAlaArgArgArg 593
QY 724 CGCCACGGCATCATCCGACGCGCACAT-----CGGCAAGG 762
DB 594 SerArgSerArgThrProThrArgArgSerArgSerArgThrProAlaArgArgGly 613
QY 763 TCTTCAAGCATCGGATA---CAACGACGAAACATCCAGACACGCGCTCTCGCTACCG 819
DB 614 ArgSerArgSerArgThrProAlaArgArgSerArgThrArgSerPro---ValArg 632
QY 820 CGAAACCGTCGAACAGTCGCCCTCTACCAAAATACAGACGGGACGCATCGACTGCA 879
DB 633 ArgArgSerArgSerArgSerProAlaArgArgSerGlyArgSerArgSerArgThrPro 652
QY 880 GAGCGTCCAAACCCGCTGATCAGCGACACCCCTGCAAA-----AGGACTCGACCGGA 933
DB 653 AlaArgArgGlyArgSerArgSerArgThrProAlaArgArgGlyArgSerArgSerArg 672
QY 934 CCGCGCAAAACCGCGGATTGCGGAGGCTGCAAGA----- 969
DB 673 ThrProAlaArgArgSerGlyArgSerArgSerArgThrProAlaArgArgGlyArgSer 692
QY 970 CGCGCTCAACAGCCGCAAAACGCTCTATCTGGTTCCACCTATTTCTGTCCTACAAA 1029
DB 693 ArgSerArgThrProArgArgGlyArgSerArgSerArgSerLeuValArgArgGlyArg 712
QY 1030 ATCGGCGACAGA-----CGCACTGGCAAACTGTGCGAGCGACGCATAGAGTTAC 1080
DB 713 SerHisSerArgThrProGlnArgArgGlyArgSerGlySerSerSerGluArgLys--- 731
QY 1081 CGTCTGACCAACTCGCTACAGCGCGACCGGCTTTCGCGC----- 1119
DB 732 -----AsnLysSerArgThrSerGlnArgArgSerArgSerAsnSerSerProGluMet 749
QY 1120 -----CGTCCATTCCGGCTAGCTCAATACCGAATAACCGAACCCTGCT 1158
DB 750 LysLysSerArgLysSerArgSerArgSerLeuSerLeuSerProArgSerLysala 769
QY 1159 CAAGCGCGCATCAAACTCTACGA-----GCTGCAACCAACCATCGCTCCCGCCAC 1212
DB 770 LysSerArgLeuSerLeuArgArgSerLeuSerLysSerProCys---ProLysGln 788
QY 1213 AAAAGACAAAGCGCTGACCGCGCAGCTCCGT-----AACCAAGCTGC----- 1253
DB 789 LysSerGlnThrProArgArgSerArgSerGlySerGlnProLysAlaLysSer 808
QY 1254 -----ATGCCAAACCTTCATTCTGTGGACGCGCA 1280
DB 809 ArgThrProProArgArgSerArgSerSerSerProProLysGlnLysSerLys 828
QY 1281 AACGCATCTTTCATCGGCTCATCAACCTCGACCCCGCTCCG----- 1322
DB 829 ThrProSerArgGlnSerHisSerSerSerProHisProLysValLysSerGlyThr 848
QY 1323 -----CACGGCTCAATACCGAAATGGCGCTGCTCATCGAAAGCCCAAAATCGCAG 1373
DB 1323 -----CACGGCTCAATACCGAAATGGCGCTGCTCATCGAAAGCCCAAAATCGCAG 1373

Db 849 ProProArgGlnGlySerIleThrSerProGlnAlaAsnGluSerValThrProGln 868
QY 1374 AACAGATGAGCGCACCTCGCGGATACACACCGAATACCCGTACCGGTACCTCG 1433
Db 869 ArgArg---SerCysPheGluSerSerProAspProGluLeuLysSerArgThrProSer 887
QY 1434 ACAACACACACCGCTCGCAATGCG-----ACGATCCCGCCA 1469
Db 888 ArgHis-----SerCysSerGlySerSerProArgValLysSerSerThrProPro 905
QY 1470 CCGGAAAAACCTACCCGA---ACGAACCGGAAGCCA----- 1502
Db 906 ArgGlnSerProSerArgSerSerProGlnProLysValLysAlaIleIleSerPro 925
QY 1503 -----AACTTGGAAACGATCGCGCAAAAATCCTATCCCTGTGCGCCATCG 1550
Db 926 ArgGlnArgSerHisSerGlySerSerProSer-----ProSer 939
QY 1551 AAGGTT 1556
Db 940 ArgVal 941
RESULT 35
ADP24102
ID ADP24102 standard; protein; 2752 AA.
XX
AC ADP24102;
XX
DT 18-NOV-2004 (first entry)
XX
DE PRO polypeptide SEQ ID NO:1280.
XX
KW PRO; antiinflammatory; antiarthritic; antirheumatic; immunosuppressive;
KW osteopathic; antidiabetic; dermatological; antipsoriatic; antiallergic;
KW antiasthmatic; hepatotropic; respiratory; gene therapy; immune system.
XX
OS Unidentified.
XX
PN WO2004041170-A2.
XX
PD 21-MAY-2004.
XX
PF 30-OCT-2003; 2003WO-US034312.
XX
PR 01-NOV-2002; 2002US-0423394P.
XX
PA (GETH) GENENTECH INC.
XX
PI Clark H, Schoenfeld J, Van Lookeren M, Williams PM, Wood WI;
PI Wu TD;
XX
DR WPI; 2004-419628/39.
DR N-PSDB; ADP24101.
XX
PT New PRO polypeptides and polynucleotides, useful for treating e.g.
PT erythematous, rheumatoid arthritis, diabetes mellitus, immune-mediated
PT renal disease, or demyelinating diseases of the central or peripheral
PT nervous system.
XX
PS Claim 7; SEQ ID NO 1280; 2940pp; English.
XX
CC The invention relates to a novel isolated nucleic acid and the PRO
CC polypeptide encoded by it. A protein of the invention has
CC antiinflammatory, antiarthritic, antirheumatic, immunosuppressive,
CC osteopathic, antidiabetic, dermatological, antipsoriatic, antiallergic,
CC antiasthmatic, hepatotropic, and respiratory activity. A polynucleotide
CC of the invention may have a use in gene therapy. The PRO polypeptide, its
CC agonist, antagonist, or antibody that specifically binds to the
CC polypeptide is useful for treating an immune related disorder such as
CC systemic lupus erythematous, rheumatoid arthritis, osteoarthritis,
CC juvenile chronic arthritis, a spondyloarthropathy, systemic sclerosis, an
CC idiopathic inflammatory myopathy, Sjogren's syndrome, systemic
CC vasculitis, sarcoidosis, autoimmune haemolytic anaemia, autoimmune

CC thrombocytopenia, thyroiditis, diabetes mellitus, immune-mediated renal
CC disease, a demyelinating disease of the central or peripheral nervous
CC system, idiopathic demyelinating polyneuropathy, Guillain-Barre syndrome,
CC a chronic inflammatory demyelinating polyneuropathy, a hepatobiliary
CC disease, infectious or autoimmune chronic active hepatitis, primary
CC biliary cirrhosis, granulomatous hepatitis, sclerosing cholangitis,
CC inflammatory bowel disease, gluten-sensitive enteropathy, Whipple's
CC disease, an autoimmune or immune-mediated skin disease, a bullous skin
CC disease, erythema multiforme, contact dermatitis, psoriasis, an allergic
CC disease, asthma, allergic rhinitis, atopic dermatitis, food
CC hypersensitivity, urticaria, an immunologic disease of the lung,
CC eosinophilic pneumonia, idiopathic pulmonary fibrosis, hypersensitivity
CC pneumonitis, a transplantation associated disease, graft rejection or
CC graft-versus-host disease. The present sequence represents a PRO protein
CC of the invention.
XX
SQ Sequence 2752 AA;

Alignment Scores:
Pred. No.: 1.87e-13 Length: 2752
Score: 303.50 Matches: 156
Percent Similarity: 37.6% Conservative: 78
Best Local Similarity: 25.1% Mismatches: 233
Query Match: 10.6% Indels: 156
DB: 8 Gaps: 29

US-10-665-990A-13 (1-1561) x ADP24102 (1-2752)
QY 43 AACACGACGCTCATTTCCCTTTTATGCTCTCTCTCTCTCATGTTCTTCTCATGTTGCC 102
Db 358 SerThrGlyProGluProProAlaProThrProLeu----- 370
QY 103 CCACCTGGAAGAACCGAGCGGAAAGCGTCATTTCAATCTCTCCAAACCTGCTCTCTGGA 162
Db 371 ---AlaGluArgHisGlyGlySerPro-----GlnProLeuAlaThrThrProLeuSer 387
QY 163 CAACATCTGCAATCCG-----GCACACCCCTCATACACAGCGCTATCCGACAT 213
Db 388 GlnGluProValAsnProProSerGluAlaSerProThrArgAspArgSerProProLys 407
QY 214 CTACCTGTCTCGACGACCC-----CCACGAGCCCTTGCCGC 249
Db 408 SerProGluLysLeuProGlnSerSerSerGluSerSerProSerPro---Gln 426
QY 250 CCGCGCCGCTTATCGAATCTGCGAACACAGCTCGATTTGCACTACTACTATTGGCG 309
Db 427 ProThrLysValSerArgHisAlaSerSerProGluSerProLysProAla 446
QY 310 CAACGACATTTCCGCGAGCTGCTGTTCAACCT----- 342
Db 447 ProGlySerHisArgGluLeuSerSerSerProThrSerLysAsnArgSerHisGlyArg 466
QY 343 -----CATGTACCTTCCGAGAACCGCGGTGCGGTACG 378
Db 467 AlaLysArgAspLysSerHisSerHisThrProSerArgArgMetGlyArgSerArgSer 486
QY 379 CTGCTGTGTGAGCAGACACACACGCGGGTGGACGATCTCTCT----- 423
Db 487 ProAlaThrAlaLysArgGlyArgSerArg---SerArgThrProThrLysArgGlyHis 505
QY 424 GTCGCGCTCGACAGCCATCCCAATATCGAAGTGGCGCTGTTCAACCCCTTCTGCTCAGC 483
Db 506 SerArgSerArgSerProGlnTrpArgSerArgSerArgSerAlaGln----- 520
QY 484 AAAATGGCGCGCACTCGGCTACTCGACCGGACTTCCCGCCCTCAACCGCGGATGCACAA 543
Db 521 ArgTpGlyArgSerArgSerProGlnArgArgGlyArgSerArgSerPro----- 537
QY 544 CAAATCTTTTACCGCGCAACCGCGCCACCATCTCGCGGCGGACCAATATCGGCGCAG 603
Db 538 -----GlnArgProGlyTrpSerArgSerArgAsnThrGlnArgArgGlyArg 553
QY 604 ATACTTCAAAGTCGGTGAGGACACCGTTTTTCGCCGACCTGGACATCTCTCGCCACCGCAG 663

```

Db 554 SerArgSerAlaArgGlyArgSerHisSerArgSerProAlaThrArgGlyArgSer 573
QY 664 CGTCCTCGCGCAAGTATCGCAGCATTCGACGGTACTGGCGCAAGCCATTCGCGCCACAA 723
Db 574 ArgSerArgThrProAlaArgGlyArgSerArgSerArgThrProAlaArgArg 593
QY 724 CGCCACGCGCATCATCCGCGCGCAACAT-----CGCAAGGG 762
Db 594 SerArgSerArgThrProThrArgArgArgSerArgSerArgThrProAlaArgGly 613
QY 763 TCTTCAAGCACTCGGATA---CAACGACGAACATCCAGACCGCTCTCGCCTACCG 819
Db 614 ArgSerArgSerArgThrProAlaArgArgSerArgThrArgSerPro---ValArg 632
QY 820 CGAAACCGTCCAGACGTCGCCCTCTACCAAAAATACAGACGGCGCATCGACTGCCA 879
Db 633 ArgArgSerArgSerArgSerProAlaArgSerGlyArgSerArgSerArgThrPro 652
QY 880 GAGCGTCCAAACCGCGCTGATCAGCAGCACCCCTGCAAA-----AGGACTCGACCGCA 933
Db 653 AlaArgGlyArgSerArgSerArgThrProAlaArgArgGlyArgSerArgSerArg 672
QY 934 CGCGCGCAACCGCGATTCGCGGAGGCTGCAAGA----- 969
Db 673 ThrProAlaArgArgSerGlyArgSerArgSerArgThrProAlaArgArgGlyArgSer 692
QY 970 CGCGCTCAACAGCCGCAAAAGCGTCTATCTGGTTTCACCCCTATTCGTCCTACAAA 1029
Db 693 ArgSerArgThrProAlaArgGlyArgSerArgSerArgSerLeuValArgGlyArg 712
QY 1030 ATCCGCGCACAGA-----CGCACTGGCAAAACTGGTCAGCAGCGCATAGACGTAC 1080
Db 713 SerHisSerArgThrProGlnArgArgGlyArgSerGlyArgSerSerGluArgLys--- 731
QY 1081 CGTCCTGACCAACTCGCTACAGCGCAGCGCGCGCTGCGCC----- 1119
Db 732 -----AsnLysSerArgThrGlnArgArgSerArgSerArgSerArgSerProGluWet 749
QY 1120 -----CGTCATTCGGTACGTCAATACCGAAACCGCGTCT 1158
Db 750 LysLysSerArgIleSerArgArgSerArgSerLeuSerArgSerProArgLysAla 769
QY 1159 CAAAGCCGCGCATCAAACTCTACGA-----GCTCAACCCCAACCATCGCTCCCGCCAC 1212
Db 770 LysSerArgLeuSerLeuArgArgSerLeuSerGlySerProCys---ProLysGln 788
QY 1213 AAAAGCAAAGCGCTGACCGCAGCTCGT-----AACCAAGCTGC----- 1253
Db 789 LysSerGlnThrProProArgArgSerArgSerGlySerGlnProLysAlaLysSer 808
QY 1254 -----ATGCCAAACCTTCATTGTGGACGGCA 1280
Db 809 ArgThrProProArgArgSerArgSerSerSerProProLysGlnLysSerLys 828
QY 1281 AACGCATCTTCATCGGCTCAATCAACTCGACCCCGCTCG----- 1322
Db 829 ThrProSerArgGlnSerHisSerSerSerProHisProLysValLysSerGlyThr 848
QY 1323 -----CACGGCTCAATACCGAATGGCGTGCATCGAAGCCCAAAATCCGAC 1373
Db 849 ProProArgGlnGlySerIleThrSerProGlnAlaAsnGluGlnSerValThrProGln 868
QY 1374 AACAGATGGAGCGCACCTCGCCGATACCAACCCGGAATACGCTACCGGTACCTCG 1433
Db 869 ArgArg---SerCysPheGluSerSerProAspProGluLeuLysSerArgThrProSer 887
QY 1434 ACAACACAACCGCTCAATGGC-----ACGATCCCGCCA 1469
Db 888 ArgHis-----SerCysSerGlySerProProArgValLysSerSerThrProPro 905
QY 1470 CCGGAAACCTACCGCA---ACGAACCCGAAGCCA----- 1502

```

906 ArgGlnSerProSerArgSerSerProGlnProLysValLysAlaIleSerPro 925

1503 -----AATTTGGAACGATCGCGCAAAATCCTATCCTGCTGCGCCATCG 1550

926 ArgGlnArgSerHisSerGlySerSerProSer-----ProSer 939

1551 AAGGTT 1556

940 ArgVal 941

RESULT 36

ID ADY91619 standard; protein; 2752 AA.

XX ADY91619;

XX 16-JUN-2005 (first entry)

XX Human prostatic cancer marker, splicing coactivator SRM300 protein.

XX tumor marker; cytostatic; prostate tumor; andrology;

XX genitourinary disease; neoplasm; immunotherapy;

XX splicing coactivator subunit SRM300.

XX Homo sapiens.

XX JP2005080524-A.

XX 31-MAR-2005.

XX 05-SEP-2003; 2003JP-00313565.

XX 05-SEP-2003; 2003JP-00313565.

XX (DOKU-) DOKURITSU GYOSEI HOJIN KAGAKU GIJUTSU SH.

XX WPI; 2005-266585/28.

XX N-PSDB; ADY91654.

XX Novel prostatic-cancer marker polypeptide, useful for diagnosing

XX prostatic-cancer or androgen independent prostatic-cancer.

XX Claim 1; SEQ ID NO 18; 43pp; Japanese.

XX The invention relates to a novel prostatic cancer marker polypeptide

XX comprising any one of 35 fully defined amino acid sequences (SEQ ID Nos:1

XX -35) as given in specification. Prostate-specific antigen (PSA) is

XX currently widely used as a prostatic tumor marker. Certain problems are

XX associated with diagnosis using PSA, however, such as false positives and

XX ambiguity between benign and malignant results. The polypeptides of the

XX invention demonstrate cytostatic activity and may be useful for

XX diagnosing and treating prostatic cancer or androgen-independent

XX prostatic cancer. The polypeptides may be utilized for immunotherapy. The

XX current sequence is that of the human prostatic cancer marker, splicing

XX coactivator subunit SRM300 protein of the invention.

XX Sequence 2752 AA;

Alignment Scores:

Pred. No.: 1.87e-13 Length: 2752

Score: 303.50 Matches: 156

Percent Similarity: 37.6% Conservative: 78

Best Local Similarity: 25.1% Mismatches: 233

Query Match: 10.6% Indels: 156

DB: 9 Gaps: 29

US-10-665-990A-13 (1-1561) x ADY91619 (1-2752)

QY 43 AACACGAGCGCTCATTCCTTTATGCGCTCTCTCTGTCATGTTCTCATGGTTGCC 102

Db 358 SerThrGlyProGluProAlaProThrProLeuLeu----- 370

QY 103 CCCACTGGAAGAACGACGAGCGGAAACCGCTCATTTCAATACTTCCAAACCTGCTCCTCTGGA 162

(GENO-) GENOME THERAPEUTICS CORP.

Rubenfield MJ, Nolling J, Deloughery C, Bush D;

WPI; 2003-615309/58.

N-PSDB; ABD01937.

Novel isolated nucleic acid encoding *Pseudomonas aeruginosa* polypeptide, useful as molecular targets for diagnostics, prophylaxis and treatment of pathological conditions resulting from bacterial infection.

Disclosure; SEQ ID NO 17112; 455pp; English.

The invention relates to *Pseudomonas aeruginosa* polypeptides and the polynucleotides encoding them. The sequences are useful in diagnosis and therapy of pathological conditions, as molecular targets for diagnostics, prophylaxis and treatment of pathological conditions resulting from a bacterial infection, for evaluating a compound, such as a polypeptide, for the ability to bind a *P. aeruginosa* nucleic acid, as components of effective antibacterial targets, as targets for antibacterial drugs, including anti-*P. aeruginosa* drugs, as templates for recombinant production of *P. aeruginosa*-derived peptides or polypeptides, as target components for diagnosis and/or treatment of *P. aeruginosa*-caused infection, and in detection of *P. aeruginosa* sequences or other sequences of *Pseudomonas* species using biochip technology. Sequences ABO67826-ABO84396 represent *P. aeruginosa* polypeptides of the invention. Note: The sequence data for this patent did not form part of the printed specification but was obtained in electronic format from USPTO at seqdata.uspto.gov/sequence.html

SQ Sequence 655 AA;

Alignment Scores:

Pred. No.:	1.44e-13	Length:	655
Score:	303.00	Matches:	155
Percent Similarity:	33.5%	Conservative:	37
Best Local Similarity:	27.1%	Mismatches:	179
Query Match:	10.6%	Indels:	202
DB:	7	Gaps:	29

US-10-665-990A-13 (1-1561) x ABO68366 (1-655)

Qy	157	CTGGACACACATCTGCAAAAT-----CGGCACACCCCTCATACAA	198
Db	140	ProAlaGluGluProAlaGlyValArgArgGluLeuLeuProArgHisProAlaGlyGly	159
Qy	199	CGG-----GCTATCGACATCTACTGCTGCAGCACCCTCGAGCGAGCCCTTGC	246
Db	160	ArgAspProAlaGlyArgLeuArgLeuValProArgArgProProArgHis-----	177
Qy	247	CGCCGCGCGCCCTTATCGAATCTGCGACACAGCCTCGATTGCGAGCAATACATTG	306
Db	178	-----TrpArgGlyLeuArgValArgAlaValProArgHisAspLeuGlyArgAla	195
Qy	307	CGCAACAGACATTTTCGCGAGGCTGCTTCAACCTCATGTACCTTGCAGCAACCGCG	366
Db	196	AspGluAlaArgLeuArgArgCysProAlaProGlyLeuAlaProProArgArgThrLeu	215
Qy	367	CGTGGGTAGCCCTGCTGTTGGAGCACAACACGCGCGGGTTGACAGATCTCTGCT	426
Db	216	ArgArgArg-----ProLeuHisArgProGlnGlyAlaArgProGlyArgSerAlaGly	233
Qy	427	CGCCCTCGACAGCCATCCCATATCGA-----AGTGGCGCTGTTCACCC	471
Db	234	AlaAlaGlyArgGlyLysAlaGlnArgLeuHisLeuGluProValAlaArgValArgArg	253
Qy	472	CTTCGTCTCTAGCAAAATGGCGCAGCTCGCTACCTGAC-----CGACTTCCCGCCCT	525
Db	254	LeuArgAlaValArgGlnValArgGlyAspValProGlyValArgArgArgProAlaAla	273
Qy	526	CAACCGCGCATGTCACAAATCTTTACCGCGCACAACCGCGCCACCATCTCTCGCGG	585
Db	274	GluProGluGluAla-----AspProGlyHis-----	282

Qy	586	ACGAATATCGCGCAGCAATACTTTCAAAGTCGGTGAGGACACCGTTTTTCGCGACCTCGA	645
Db	283	-----GlyHis-----ArgProGly	287
Qy	646	CATCTCGCCACCGCGCAGCGTCTGCGCGAAAGTATCGCACGACTTCGACCGCTACTGGCG	705
Db	288	ArgArgGlnArgGluValArgArgGlnProLeuSer-----ArgGlnThrAlaGly	305
Qy	706	AAGCCATTCGCG-----CCACAACGC	726
Db	306	ArgAlaArgArgProAlaProAlaAspArgAlaGlyTrpGlnGlyProGlyArg	325
Qy	727	CACGCGCATCTCCGACGCGCAACATCGCGAAGGGTCTTCAAGCACTCGGATACAAAGA	786
Db	326	ArgArgHisProValValLeuHisLeuProGly-----LeuArg	339
Qy	787	CGAAACATCCAGACACACGCGTCTCTGCGCTACCGGAAACCGTCGAACAGTCCGCCCTCTA	846
Db	340	ArgGlyValProAspAspArgAlaArgArgHisArgHis-----AlaProPro	358
Qy	847	CCAAAAATACAGAGCGGACGCATCGACTGCGCAGCGT-----CCAAACCCG	894
Db	359	Pro-----AspProGlyGluArgArgHisAlaGlnGlnGlyArgArgGlyValGlyGln	376
Qy	895	CCTGATCAGCGACACCCCTGCAAAAGGACTCGACCG-----	930
Db	377	ProAspArgHisArgGlnProArgArgPheArgProArgArgProValGluLeuGlyGly	396
Qy	931	CGACCGCGCAACCGCGATTGCGGGAGGCTCGAAGACGC-----	972
Db	397	ArgProGluProAlaAlaAspGluArgGluGluArgArgAlaValLeuGlyArg	416
Qy	973	-----GCTCAACACGCGCGAAAAAGCGTCTA-----	999
Db	417	ArgTrpArgLeuArgHisAlaGlnProAlaHisProAlaArgLeuArgGluAspProGln	436
Qy	1000	-----TCTGGTTTTTCCCTATTTTCGTCCTTACAAAATCCCGGCACAGCGACTCGC	1050
Db	437	GlySerArgGlyGlyLeuArgArgAlaArgProGlyGlyAlaArgGlnArgArgGly	456
Qy	1051	AAAACTGGTCGAGCAGCG-----	1068
Db	457	---AlaProAlaGlyArgArgGlyTyLeuProGluProGlyGlnAlaGlnTyArgAsp	475
Qy	1069	-----CATAGAGTTACCGTCTCTGACCACTCGCTACAGGC	1104
Db	476	ProGlyGluIleProLeuHisProHisArgGlnLeuArgProAlaGlnLeuProCysAla	495
Qy	1105	GACCGAGTTTGGCGCGCTCCATTCGCGTACGTCAAATACCGAAAAACCGCTGCTCAAAGC	1164
Db	496	GluGluArgIleArgArgPro-----ArgArgArgLeuProGlyThrAlaProGlnHis	513
Qy	1165	CGGCATCAAACTCTACGAGCTGCAACCCCAACATGCGCTGCCGCCCAAAAGACAAAG	1224
Db	514	LeuHisArg-----ArgThrAlaAlaArg-----ArgHisAlaAlaThrGly	527
Qy	1225	CCTGACCGGAGCTCGCTTAACCGCTGATGCCAAAACCTTATTGTGGACGGCAACG	1284
Db	528	ProGlnGln-----GlyArgGlnArg	534
Qy	1285	CATCTTCATCGGCTCATTCAACTTCGACCCCGTTCGCGCGCTCAATACCGAAATCGG	1344
Db	535	HisLeuProArgProLeuLeuPro-----	542
Qy	1345	CGTCGTATCGAAGCCCAAAATTCGCAACACAGATGGAGCGCACCTCGCCGATACAC	1404
Db	543	-----GlyProLeuGlnArgArgValArgGlyAlaAlaGlnArgAlaAlaArg	558
Qy	1405	ACCGCA-----ATAGCGCTACCG-----	1422
Db	559	ThrArgHisArgGlyGlnGlyAsnAlaThrLeuArgLeuProLeuAlaLeuLeuArgArg	578

```
QY 1423 -----CGTTACCTCGACAAACACACCGCTGCAATGCGACGA 1461
Db 579 ArgTrpArgArgAlaAspHisArgHisProArgGlnAlaThrAspSerArgHisAlaHis 598
QY 1462 TCCCGCCGCCGAAACCTACCCGAAACGACGACCGCAACCTTTGGAAACGCAATCGC 1521
Db 599 GlyArgHisProArgAsp-----Arg 605

QY 1522 CCACAA-----AATCTATCCCTGCTGCCCCATCGA 1551
Db 606 ArgArgAlaGlyGlyArgArgLeuProAlaValHisArg 618

RESULT 38
ABO74950
ID ABO74950 standard; protein; 618 AA.
XX
XX ABO74950;
XX
DT 29-JUL-2004 (first entry)
XX
XX Pseudomonas aeruginosa polypeptide #7125.
DE
XX Bacterial infection; Pseudomonas aeruginosa infection; antibacterial.
KW
XX Pseudomonas aeruginosa.
OS
XX US6551795-B1.
PN
XX 22-APR-2003.
PD
XX
XX 18-FEB-1999; 99US-00252991.
PF
XX
XX 18-FEB-1998; 98US-0074788P.
PR
XX 27-JUL-1998; 98US-0094190P.
PX
XX (GENO-) GENOME THERAPEUTICS CORP.
PA
XX
XX Rubenfield MJ, Nolling J, Deloughery C, Bush D;
PI
XX WPI; 2003-615309/58.
DR
XX N-PSDB; ABD08521.
XX
XX Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide,
PT useful as molecular targets for diagnostics, prophylaxis and treatment of
PT pathological conditions resulting from bacterial infection.
XX
XX Disclosure; SEQ ID NO 23696; 455pp; English.
XX
XX The invention relates to Pseudomonas aeruginosa polypeptides and the
CC polynucleotides encoding them. The sequences are useful in diagnosis and
CC therapy of pathological conditions, as molecular targets for diagnostics,
CC prophylaxis and treatment of pathological conditions resulting from a
CC bacterial infection, for evaluating a compound, such as a polypeptide,
CC for the ability to bind a P. aeruginosa nucleic acid, as components of
CC effective antibacterial targets, as targets for antibacterial drugs,
CC including anti-P. aeruginosa drugs, as templates for recombinant
CC production of P. aeruginosa-derived peptides or polypeptides, as target
CC components for diagnosis and/or treatment of P. aeruginosa-caused
CC infection, and in detection of P. aeruginosa sequences or other sequences
CC of Pseudomonas species using biotech technology. Sequences ABO67826-
CC ABO84396 represent P. aeruginosa polypeptides of the invention. Note: The
CC sequence data for this patent did not form part of the printed
CC specification but was obtained in electronic format from USPTO at
CC seqdata.uspto.gov/sequence.html
XX
XX Sequence 618 AA;

Alignment Scores:
Pred. No.: 1.68e-13 Length: 618
Score: 302.00 Matches: 173
Percent Similarity: 34.8% Conservative: 35
Best Local Similarity: 29.0% Mismatches: 193
Query Match: 10.6% Indels: 196
```

```
DB: 7 Gaps: 38
US-10-665-990A-13 (1-1561) x ABO74950 (1-618)
QY 169 CTGCAAAATCGGACACCCCTCATACAAACGGGCTATCCGACATCTACCTGCTCGACGA 228
Db 4 ProaspHisProArgHisProArg---ArgArgGlnLeuAlaGluLeuArgIleArgArg 22
QY 229 -----CCCCACGAAGCCCTTTCGCGCCCGCCGCTTATCGAATCTGCGCAACA 279
Db 23 ArgThrAlaProGlySerAspArgGlnGlnProArgArgThrProArgIleArgArgArg 42
QY 280 CAGCCTCGATTGCAATACTACATTGTCGCGCAACGACAT----- 318
Db 43 -----HisGlnGlyGlnSerHisArgProAlaHisGlnGlyArg 55
QY 319 -----TTCCGGCAGGCTGCTTCAACCTCAT 345
Db 56 GlnArgGlnProGlyAlaProAlaAlaMetGlyLeuArgArgAla-----TrpProAla 73
QY 346 GTACCTTCGCGACGAACCGCGTGCCTGCTGTTGGAGACGACAAACACGCG 405
Db 74 AlaProCysGlyArgArgArgProAspThr-----GlnLeuArgLeu 88
QY 406 CGGGTTGACGATCTCCTGCTGCGCCTCGACAGCCATCCCAATATCGAAGTCGCGCTGTT 465
Db 89 ArgProGlnArgGlnSerGlyArgArgAsnGlnPro-----AlaProVal 103
QY 466 CAACCCCTT-----ATCCTTACCGCGCAACCG----- 567
Db 144 ArgCysHisHisProLeuArgIleArgArgProArgGlnProAspProThrGlyGlnPro 163
QY 568 -----CGCCACCACTACTCGGACGCAATATCGCGACGAATACTTCAAAAGTCGGTGA 621
Db 184 GlyGlnArgHisHisHisLeuArgAlaArgArgArg-----GlnArgHisPro 180
QY 622 GGACACCGTTTTTCGCGACCTCGACATCTCGCCACCGCGACGCTGTCGGCGAAGTATC 681
Db 181 ProHisArgArgSerArgCysGlyHis---ArgValSerLeuArgArgProGlnSerAla 199
QY 682 GCACGACTTCGACCGCTACTGGGCAAGCCATTCGCCCCACACGCGCATCATCCG 741
Db 200 GlyArg---AlaProLeuAlaGluArgPro---GluProArgArgThrVal-----Pro 215
QY 742 CAGCGGCAACATCGGCAAGGCTTCTCAAGCACTCGGATACAA---CGACGAAACATCCAG 798
Db 216 LeuArgProHisArgArgArgGlnGlnGlyHisArgProProGlyArgHisArgTrpArg 235
QY 799 ACACGCGCTCTCGCTACCGCGAAACCGTCGAAACAGTCGCGCCCTCTACCAAAAAATACA 858
Db 236 SerArgGlnProGlyValProLeuArgArg----- 245
QY 859 GACGGGACGTCGATCGATGGCAGCGTCCAAACCGCCCTGATCAGCAGACACCCCTGCAAA 918
Db 246 -----AlaArgGlnProGlyArgAlaGlyAlaGlnHisPro----- 257
QY 919 AGGACTCGACCGCGACCGCCGCAACCGCATTCGCGGAG---GCTGCAAGACGCGCT 975
Db 258 -----ProArgProAlaAspProAlaArgProArgAspLeuProLeuArgArgGly 274
QY 976 CAACACGCGCGAAAAACGGTCTATCTGTTTACCCCTTATTCCTTACAAAATCCCG 1035
Db 275 GlnProThrAlaGlyAspArgLeu-----ProLeuArgProArg 287
QY 1036 CAC-----AGCGCACTGCGCAA-----ACTGTCGACGACCG 1068
```

Db 288 HisArgLeuProAlaGlnArgArgProGlyArgGlnArgAspProGlySerGlyArg 307
QY 1069 CATACAGCTTACCGTCTCT-----GACCACTCGGTACA----- 1101
Db 308 GlnGlyAlaGluHisProGlyArgAlaAspArgLeuProAlaLeuArgProAlaAlaAla 327
QY 1102 -----GGCGACCGAGTTCGCGCTCCGCTCATTCCGCTACGTCAA---ATACCGAAA 1149
Db 328 ProAspLeuGlyGlnArgHisSerGlnPro---ArgValArgProGlyLeuProThr 346
QY 1150 ACCGCTGTCTCAAGCCGG-----CATCAAACTCTACGAGCTGCAACCCAA----- 1194
Db 347 AlaAlaAlaGluGlyArgProAlaGluArgLeuProAlaArgCysGlnTrpGlnTrp 366
QY 1195 -----CGATCCGCTCCCGCCACAAA 1215
Db 367 ProAlaAlaProGlnProLeuGlyHisProGlyLeuProValArgProAlaGlyPro 386
QY 1216 AGACAAAGGCTGACCGCAGCTCCGTAACACGAGCTGCATGCCAA---AACCTTCATTGT 1272
Db 387 -----ProAspArgGlyThrArgArgProGlyArgAlaGlnLeuCysLeuArgArg 403
QY 1273 GGACGGCAAAACGCAAT-----CTTCATCGGCTCATTCAA 1305
Db 404 GlyArgGlnProHisProThrGlnArgGlnProArgLeuArgArgHisArgGln---Gln 422
QY 1306 CTGACACCGCTTCGCGCAGCTCAATAC-----CGAATGGCGTGGTCAATCGA 1356
Db 423 ProGlyLeuProValArgAlaArgGlnGlnProIleAspArgHisArgAlaSerGly 442
QY 1357 AAGCCCCAAATCGCAGAACAGATGGAGCGCACCTCGCCATACACACCGAATACGC 1416
Db 443 AspGlnArgArgArgGlnProHisProGlyProArgArgAlaGlnThrGlyLeuArg 462
QY 1417 CTACCGCTTACCTTCGCAACACACACCGCTCGCAATGGCA----- 1458
Db 463 ArgProGlyProProAlaAlaArgGlnProArgArgProAlaGlyArgArgIleProLeu 482
QY 1459 -----CGATCCGCGCACCGGAAA-----AACCTACCC--- 1485
Db 483 GlnArgProArgProAlaHisArgGlnAlaHisProArgGluArgHisLeuProLeu 502
QY 1486 -----GAACGAACCGAAGCCAACTTTGGAACGCAT 1518
Db 503 TrpSerArgArgProValAlaArgSerArgThrArgArgGlnArgProGluThrAla 522
QY 1519 CGCCGCAAAATCCTATC-----CCTGCTGCC-----CATCGAAGG 1554
Db 523 SerProValLeuProValAlaGlyGlnProAlaAlaGlyHisHisArgArg 539

RESULT 39
ADJ70425
ID ADJ70425 standard; protein; 2263 AA.
XX AC ADJ70425;
XX DT 06-MAY-2004 (first entry)
XX DE Human heat mitochondrial protein as a therapeutic target SeqID2231.
XX KW mitochondrial; human; screening assay; diabetes mellitus;
KW Huntington's disease; osteoarthritis;
KW Leber's hereditary optic neuropathy; LHON;
KW mitochondrial encephalopathy lactic acidosis and stroke; MELAS;
KW myoclonic epilepsy ragged red fibre syndrome; MERRF; cancer;
KW neuroprotective; neurotropic; antidiabetic; anticonvulsant; antiarthritic;
KW osteopathic; ophthalmological; cytostatic.
XX OS Homo sapiens.
XX PN WO2003087768-A2.
XX XX

PD 23-OCT-2003.
XX 04-APR-2003; 2003WO-US010870.
XX 12-APR-2002; 2002US-0372843P.
PR 17-JUN-2002; 2002US-0389987P.
PR 20-SEP-2002; 2002US-0412418P.
XX (MITO-) MITOKOR.
PA Ghosh SS, Fahy ED, Zhang B, Gibson BW, Taylor SW, Glenn GM;
PI Warnock DB;
XX WPI; 2003-845369/78.
XX Identifying a mitochondrial target for drug screening assays and for
PT treating diseases associated with altered mitochondrial function,
PT comprises detecting a modified polypeptide in a sample and correlating
PT with the disease.
XX Claim 1; SEQ ID NO 2231; 180pp; English.
XX This invention relates to novel mitochondrial targets that can be used
CC for therapeutic intervention in treating a disease associated with
CC altered mitochondrial function. Specifically, it refers to a method for
CC identifying proteins of the human heart mitochondrial proteome that are
CC useful for drug screening assays, as well as therapeutic targets. The
CC present invention describes a method for identifying such proteins that
CC can be used in the treatment of various diseases associated with altered
CC mitochondrial function including diabetes mellitus, Huntington's disease,
CC osteoarthritis, Leber's hereditary optic neuropathy (LHON), mitochondrial
CC encephalopathy lactic acidosis and stroke (MELAS), myoclonic epilepsy
CC ragged red fibre syndrome (MERRF) or cancer. Accordingly, these
CC compositions have neuroprotective, neurotropic, antidiabetic,
CC anticonvulsant, antiarthritic, osteopathic, ophthalmological and
CC cytostatic activities. This polypeptide sequence is a human heart
CC mitochondrial protein of the invention.
XX SQ Sequence 2263 AA;
Alignment Scores:
Pred. No.: 2,5e-13 Length: 2263
Score: 301.50 Matches: 154
Percent Similarity: 37.5% Conservative: 79
Best Local Similarity: 24.8% Mismatches: 234
Query Match: 10.6% Indels: 156
DB: 7 Gaps: 28
US-10-665-990A-13 (1-1561) x ADJ70425 (1-2263)
QY 43 AACACGAGCTCATTTCCCTTTTATGCTCTCTCTGTCATGTTCTTCATGTTGCC 102
Db 358 SerThrGlyProGluProProAlaProThrProLeuLeu----- 370
QY 103 CCCACTGGAAGACGAGCGGAAAGCCGTCATTTCCTCAAACTGCTCTCTGGA 162
Db 371 ---AlaGluArgHisGlySerPro-----GlnProLeuAlaThrProLeuSer 387
QY 163 CAACATCCTCGAAATCCG-----GCACCCCTCATATAACACCGGCTATCCGACAT 213
Db 388 GlnGluProValAsnProProSerGluAlaSerProThrArgAspArgSerProProlys 407
QY 214 CTACCTGCTCGACGACCC-----CCAGAACGCTTGGCCG 249
Db 408 SerProGluLysLeuProGlnSerSerSerSerSerSerSerProProSerPro---Gln 426
QY 250 CGCGCGCGCTTATCGAATCTGCCGACACAGCTCGATTGCAATCTACTACTTTGGCG 309
Db 427 ProThrLysValSerArgHisAlaSerSerSerProGluSerProAlaProAla 446
QY 310 CAACGACATTTCCGCGAGGCTGCTGTTCAACCT----- 342

Db 447 ProGlySerHisArgGluIleSerSerSerProThrSerLysAsnArgSerHisGlyArg 466
QY 343 -----CATGTACTTCCCGAGAACCGCGGTGCGGTACG 378
Db 467 AlaLysArgAspLysSerHisSerHisThrProSerArgArgMetGlyArgSerArgSer 486
QY 379 CCGTGTGTGGACGACAAACACACCGCGGGTGTGACGATCTCCT----- 423
Db 487 ProAlaThrAlaLysArgGlyArgSerArg---SerArgThrProThrLysArgGlyHis 505
QY 424 GCTCCCTCCGACAGCCATCCCAATATCGAAGTGGCGCTGTTCACACCCCTTCGTCTCCTACG 483
Db 506 SerArgSerArgSerProGlnTrpArgArgSerArgSerAlaGln----- 520
QY 484 CAAATGGCGCGCACTCGGTACTACCGCACTTCCCGCCCTCAACCCCGCATGCACAA 543
Db 521 ArgTrpGlyArgSerArgSerProGlnArgArgGlyArgSerArgSerPro----- 537
QY 544 CAAATCCTTTACCGCGGCAACCGCGCCACCATCTACCGGCGACCAATATCGGCGAGA 603
Db 538 -----GlnArgProGlyTrpSerArgSerArgAsnThrGlnArgArgGlyArg 553
QY 604 ATACTTCAAAGTCGTGAGCACACCGTTTTCGCCGACCTGGACATCTCCGCCACCGGACG 663
Db 554 SerArgSerAlaArgArgGlyArgSerHisSerArgSerProAlaThrArgGlyArgSer 573
QY 664 COTCGTGGCGAAGTATCGCAGCTTCGACCGTACTGGCGCAACCCATTCGCGCCACAA 723
Db 574 ArgSerArgThrProAlaArgArgGlyArgSerArgSerArgThrProAlaArgArg 593
QY 724 CGCCAGCGCATCATCCGACGGCGACAT-----CGGCAAGG 762
Db 594 SerArgSerArgThrProThrArgArgArgSerArgSerArgThrProAlaArgArgGly 613
QY 763 TCTTCAAGCACTCGGATA---CAACAGCAAAATCCAGACACGCGCTCTCGGTACCG 819
Db 614 ArgSerArgSerArgThrProAlaArgArgArgSerArgSerArgThrPro---ValArg 632
QY 820 CGAAACCGTCTGAACAGTCCCTCTACCAAAATACAGCGGACCGCATCTGCGCA 879
Db 633 ArgArgSerArgSerProAlaArgArgSerGlyArgSerArgSerArgThrPro 652
QY 880 GAGCCTCCAAACCCGCTGATCAGGACACCCCTGCMAA-----AGGACTGACCGGCA 933
Db 653 AlaArgArgGlyArgSerArgSerArgThrProAlaArgArgGlyArgSerArgSerArg 672
QY 934 CGCGCGCAACCGCGGATTCGCGGAGGCTGCAAGA----- 969
Db 673 ThrProAlaArgArgSerGlyArgSerArgSerArgThrProAlaArgArgGlyArgSer 692
QY 970 CGCGTCAAAACGCGCAAAAGCGTCTATCTGTTTACCCCTATTCCTTTCCTCCACAA 1029
Db 693 ArgSerArgThrProArgArgGlyArgSerArgSerArgSerLeuValArgArgGlyArg 712
QY 1030 ATCCCGCACAGA-----CGCACTGGCAAACTGTGTGACGAGCGCATAGAGTTAC 1080
Db 713 SerHisSerArgThrProGlnArgArgGlyArgSerGlySerSerGluArgLys--- 731
QY 1081 COTCCTGACCAACTCGCTACAGCGACCGACGCTTCGCGG----- 1119
Db 732 -----AsnLysSerArgThrSerGlnArgArgSerArgSerAsnSerSerProGluMet 749
QY 1120 -----CGTCCATTCCGGTACGTCATTCGTTTACCCCTATTCCTTTCCTCCACAA 1158
Db 750 LysLysSerArgIleSerSerArgSerArgSerArgSerLeuSerSerProArgSerLysAla 769
QY 1159 CAAACCGCGCATCAACTCTAGA-----GCTGCAACCCCAACCATGCGTCCCGCCAC 1212
Db 770 LysSerArgLeuSerLeuArgArgSerLeuSerGlySerSerProCys---ProLysGln 788
QY 1213 AAAAGACAAAGCCCTGACCGGAGCTCCGTAACCA----- 1247
Db 789 LysSerGlnThrProProArgArgSerArgSerGlySerSerGlnThrLysAlaLysSer 808

QY 1248 -----CCCTGCATGCCAAAACCTTCAATTGTGGACGCA 1280
Db 809 ArgThrProProArgArgSerArgSerSerSerProProProLysGlnLysSerLys 828
QY 1281 AACGCATCTTCATCGGCTCATTAACCTCGACCCCGTTCG----- 1322
Db 829 ThrProSerArgGlnSerHisSerSerSerProHisProLysValLysSerGlyThr 848
QY 1323 -----CACGGCTCAATACCGAAATGGCGCTGTCATCGAAAGCCCAAAATCGCAG 1373
Db 849 ProProArgGlnGlySerIleThrSerProGlnAlaAsnGluGlnSerValThrProGln 868
QY 1374 AACAGATGAGGAGCACCTCCCGCATACACACCGAATACGCTACCGCGTTACCTCG 1433
Db 869 ArgArg---SerCysPheGluSerSerProAspProGluLeuLysSerArgThrProSer 887
QY 1434 ACAACACACCGCTGCAATGGC-----ACGATCCCGCA 1469
Db 888 ArgHis-----SerCysSerGlySerSerProProArgValLysSerSerThrProPro 905
QY 1470 CCGCAAAACCTTACCCGA---ACGAACCCGAAGCCA----- 1502
Db 906 ArgGlnSerProSerArgSerSerSerProGlnProLysValLysAlaIleSerPro 925
QY 1503 -----AACTTTGAAACGATCGCGCGCAAAATCTATCCTCGTGCCTCG 1550
Db 926 ArgGlnArgSerHisSerGlySerSerProSer-----ProSer 939
QY 1551 AAGGTT 1556
Db 940 ArgVal 941
RESULT 40
ABO79612
ID ABO79612 standard; protein; 618 AA.
XX ABO79612;
AC ABO79612;
XX 29-JUL-2004 (first entry)
XX Pseudomonas aeruginosa polypeptide #11787.
XX Bacterial infection; Pseudomonas aeruginosa infection; antibacterial.
XX Pseudomonas aeruginosa.
OS US6551795-B1.
PN 22-APR-2003.
PD 18-FEB-1999; 99US-00252991.
XX 18-FEB-1998; 98US-0074788P.
PR 27-JUL-1998; 98US-0094190P.
XX (GENO-) GENOME THERAPEUTICS CORP.
XX Rubenfield MJ, Nolling J, Deloughery C, Bush D;
XX WPI: 2003-615309/58.
DR N-PSDB; ABD13183.
XX Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide, useful as molecular targets for diagnostics, prophylaxis and treatment of pathological conditions resulting from bacterial infection.
XX Disclosure; SEQ ID NO 28358; 455pp; English.
XX The invention relates to Pseudomonas aeruginosa polypeptides and the polynucleotides encoding them. The sequences are useful in diagnosis and therapy of pathological conditions, as molecular targets for diagnostics, prophylaxis and treatment of pathological conditions resulting from a

CC bacterial infection, for evaluating a compound, such as a polypeptide, for the ability to bind a P. aeruginosa nucleic acid, as components of effective antibacterial targets, as targets for antibacterial drugs, including anti-P. aeruginosa drugs, as templates for recombinant production of P. aeruginosa-derived peptides or polypeptides, as target components for diagnosis and/or treatment of P. aeruginosa-caused infection, and in detection of P. aeruginosa sequences or other sequences of Pseudomonas species using biophysics technology. Sequences AB067826-AB084396 represent P. aeruginosa polypeptides of the invention. Note: The sequence data for this patent did not form part of the printed specification but was obtained in electronic format from USPTO at seqdata.uspto.gov/sequence.html

XX
SQ Sequence 618 AA;

Alignment Scores:

Pred. No.:	1-99e-13	Length:	618
Score:	301.00	Matches:	175
Percent Similarity:	34.6%	Conservative:	31
Best Local Similarity:	29.4%	Mismatches:	196
Query Match:	10.6%	Indels:	194
DB:	7	Gaps:	37

US-10-665-990A-13 (1-1561) x AB079612 (1-618)

QY	169	CCTGCAATTCGGGACACCCCTCATACAAACGGGCTATCCGACATCTACCTGCTCGACGA	228
DB	4	ProAspHisProArgHisProTrp--ArgArgGlnLeuAlaGluLeuArgIleArgArg	22
QY	229	-----CCCCACGAGCCCTTCGCGCCCGCGCCCTTCGTAATTCGCGAACA	279
DB	23	ArgThrAlaProGlyGlyAspArgGlnGlnProArgArgThrProArgIleArgArg	42
QY	280	CAGCTCGATTGTCATCTACATTGGCGAACACAT-----	318
DB	43	-----HisGlnGlyGlnSerHisArgProAlaHisGlnGlyArg	55
QY	319	-----TTCCGGCAGGCTGCTGTTCAACCTCAT	345
DB	56	GlnArgGlnProGlyAlaProAlaAlaMetGlyLeuArgAla-----TrpProAla	73
QY	346	GTACCTTCGGGACGAGCGGGTGGCGGTAGCGCTGTTGGACGACAAACACGCG	405
DB	74	AlaProCysGlyArgArgArgProAspThr-----GlnLeuArgLeu	88
QY	406	CGGTTGGACCATCTCTGCTCGCCCTCGACGACATCCCAATATCGAAGTGGCGCTGT	465
DB	89	ArgProGlnArgGlnSerGlyArgArgAsnGlnPro-----AlaProVal	103
QY	466	CAACCCCTT-----CGTCTTACGCAAAATGGCGGC	495
DB	104	ArgProGlnProGlyLeuArgArgProArgProAlaGlyArgProGluArgSerProArg	123
QY	496	ACTGGCTACTGACCGACTTCCCGCGCTCAACCGCGCATGCACAAACAA-----	546
DB	124	ArgGlnAspProThrArgLeuArgArgProGlyGlnProHisArgGlyGlnGlyProAla	143
QY	547	-----ATCCTTTTACCGCGCAACCG-----	567
DB	144	ArgArgHisHisProLeuArgIleArgArgProArgGlnProAspProThrGlyGlnPro	163
QY	568	-----CGCCACCATCTCGCGGACGCAATATCGCGACGAATCTTCAAGTCGGTGA	621
DB	164	GlyGlnArgHisHisLeuArgAlaArgArgArg-----GlnArgHisPro	180
QY	622	GGACACGTTTTTCGGACCTGGACATCTCGCACCGCGAGCGTGTGGCGAAGTATC	681
DB	181	ProHisArgArgSerArgCysGlyHis---ArgValSerLeuArgArgProGlnSerAla	199
QY	682	GCAGCACTTCGACCGCTACTGGCAAGCCATTCGCCCCACACGCCGCGCATCATCCG	741
DB	200	GlyArg---AlaProLeuAlaGluArgPro---GluProArgArgThrVal-----Pro	215

QY	742	CAGCGCAACATCGGCAAGGGTCTTCAAGCACTCGGATACACGACCAACATCCAGACA	801
DB	216	LeuArgProHisArg-----ArgArgGlnProGlyHis	226
QY	802	CGCGTCTCTGGCTACCGCAACCGTCAACAGTCGCCCTCTACCAAAATACACAGAC	861
DB	227	ArgProProGly-----ArgHisArg-----CysProArgGlnProGly	240
QY	862	GGGACGATCTGCTGGGAGAGCGTCCAAACCGCTGATCAGCGACACCCCTGCAAAAGG	921
DB	241	ValProLeuArgArgAlaArgGlnProGlyArgAlaGlyThrGlnHisPro-----	257
QY	922	ACTGACCGCGACCGCGCAACCGCGATTGCGGGAG-----GCTGCAAGCGCGCTCAA	978
DB	258	-----ProArgProAlaAspProAlaArgProArgAspLeuProLeuArgArgGlyGln	275
QY	979	ACAGCCGAAAAAGGCTCTATCTGTTTCCACCTATTTCGCTACAAATACCGGCAC	1038
DB	276	ProThrAlaGlyAspArgLeu-----ProLeuArgProArgHis	288
QY	1039	-----AGACGCACTGGCAAA-----ACTGGTCAGGACGGCAT	1071
DB	289	ArgLeuProAlaGlnArgArgArgProGlyArgGlnArgAspProGlySerGlyArgGln	308
QY	1072	AGCGTTACCGTCT-----GACCAACTCGCTACA-----	1101
DB	309	GlyAlaGluHisProGlyTrpAlaAspArgLeuProAlaLeuArgProAlaAlaPro	328
QY	1102	-----GGCGACGACGTTGCCGCTCATTCGCGTACGCTCAA-----ATACGAAAAACC	1152
DB	329	AspLeuGlyGlnArgHisHisSerGlnPro--ArgValArgProGlyLeuProThrAla	347
QY	1153	GCTCTCAAGCCGG-----CATCAAACTCTAGAGCTGCAACCCAA-----	1194
DB	348	AlaAlaGluGlyArgProLeuAlaGluArgLeuProAlaCysGlnTrpGlnTrpPro	367
QY	1195	-----CCATCGCTCCCGCCGCAAAAAA-----	1218
DB	368	AlaAlaProProGlnProLeuGlyHisProGlyLeuProValArgProAlaGlyPro---	386
QY	1219	CAAGGCTGACCGGACGCTCCGTAACCGCTCGATGCCAA-----AACCTTCATTGTGGA	1275
DB	387	-----ProAspArgGlyThrArgArgProGlyArgAlaGlnLeuArgLeuArgGly	404
QY	1276	CGCAACACGCT-----CTTCATCGCTCATTCACCT	1308
DB	405	ArgGlnProHisProThrGlnArgGlnProArgLeuArgArgHisArgGln---GlnPro	423
QY	1309	CGACCCCGTTCCGCGCGCTCAATAC-----CGAAATGGCGCTCGTCATCGAAAG	1359
DB	424	GlyLeuProValArgAlaArgGlnGlnProIleAspArgHisArgArgAlaSerGlyAsp	443
QY	1360	CCCCAAATCGCAGAACAGATGGAGCGACCTCGCGATACCCACCCGGAATACGCTTA	1419
DB	444	GlnArgArgArgGlnProHisProGlyProArgArgAlaGlnThrGlyLeuArgArg	463
QY	1420	CGCGTTACCTCGACAAACACACCGCTCGCAATGCA-----	1458
DB	464	ProGlyProProAlaLysArgGlnProArgProAlaGlyArgArgIleProLeuGln	483
QY	1459	-----CGA-----	1461
DB	484	ArgProArgProAlaHisArgGlnAlaHisProArgGluHisHisLeuProLeuArg	503
QY	1462	TCCCGCACCGCAAAACCTTACCGGAACCGAACCGCAACCTTTGGAAACCGCATCGC	1521
DB	504	SerArgArgProValAlaArgSerArgThrArgArgGlnArgProGluThrAlaSer	523
QY	1522	CGCAAAATCCTATC-----CCTGCTGCC-----CATCGAAGG	1554
DB	524	ProValLeuProValAlaGlyGlnProAlaAlaGlyHisHisArgArg	539

Search completed: May 2, 2006, 05:12:31
Job time : 390 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: May 2, 2006, 05:52:55 ; Search time 49.2 Seconds
(without alignments)
2651.348 Million cell updates/sec

Title: US-10-665-990A-13

Perfect score: 2852

Sequence: 1 caaatacaggaatgcgcgt.....tgcccatcgaaggtttatta 1561

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 3735138

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+n2p.model -DEV=xlp
-Q=/abss/ABSSWEB_spo01/US10665990/runat_01052006_112001_9763/app_query.fasta_1
-DB=PublishedApplications_AA_Main -QFMT=fastan -SUFFIX=rapbm -MINMATCH=0.1
-LOOPEL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=40 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0
-MAXLEN=2000000000 -HOST=abss02p
-USER=US10665990 @CGN_1_1_307@runat_01052006_112001_9763 -NCPU=6 -ICPU=3
-NO_WMAP -NEG_SCORES=0 -WAIT -DSFLOCK=100 -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications AA_Main:

1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/US11_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2689	94.3	525	5	US-10-665-990A-14
2	2681	94.0	525	5	US-10-665-990A-16
3	2677	93.9	525	5	US-10-665-990A-18
4	2630	92.2	525	5	US-10-066-551-4
5	2630	92.2	525	5	US-10-665-990A-4
6	2630	92.2	525	5	US-10-665-990A-20
7	632	22.2	502	4	US-10-335-977-9085
8	632	22.2	502	4	US-10-335-977-9086
9	590	20.7	428	3	US-09-881-752A-356
10	359	12.6	224	4	US-10-335-977-9084
11	336.5	11.8	5179	3	US-09-922-217-1068

12	336.5	11.8	5179	3	US-09-833-263-1068	Sequence 1068, Ap
13	336.5	11.8	5179	4	US-10-025-380-1068	Sequence 1068, Ap
14	336.5	11.8	5179	4	US-10-734-584-121	Sequence 121, App
15	323.5	11.3	1367	3	US-09-801-368-108	Sequence 108, App
16	310	10.6	563	4	US-10-437-963-198755	Sequence 198755,
17	303.5	10.6	2296	5	US-10-696-909A-46	Sequence 46, Appl
18	303.5	10.6	2752	5	US-10-696-909A-44	Sequence 44, Appl
19	301.5	10.6	2263	4	US-10-408-765A-2231	Sequence 2231, Ap
20	300	10.5	528	3	US-09-840-746-20	Sequence 20, Appl
21	298	10.4	497	4	US-10-437-963-125004	Sequence 125004,
22	297.5	10.4	19723	4	US-10-084-846A-5	Sequence 5, Appli
23	296	10.4	486	4	US-10-418-861B-59	Sequence 59, Appl
24	293.5	10.3	866	5	US-10-626-832-86	Sequence 86, Appl
25	293	10.3	465	4	US-10-437-963-174113	Sequence 174113,
26	292.5	10.3	533	4	US-10-437-963-176617	Sequence 176617,
27	289.5	10.2	495	4	US-10-724-972A-3854	Sequence 3854, Ap
28	286.5	10.0	493	3	US-09-828-523A-22	Sequence 22, Appl
29	286.5	10.0	493	3	US-09-966-521-18	Sequence 18, Appl
30	286.5	10.0	493	4	US-10-429-09A-18	Sequence 18, Appl
31	286.5	10.0	502	3	US-09-828-523A-90	Sequence 90, Appl
32	286.5	10.0	502	3	US-09-966-521-84	Sequence 84, Appl
33	286.5	10.0	502	4	US-10-429-09A-84	Sequence 84, Appl
34	283	9.9	19695	4	US-10-084-846A-3	Sequence 41094, A
35	281	9.9	2284	6	US-11-097-143-41094	Sequence 41094, A
36	280.5	9.8	820	5	US-10-684-422-44	Sequence 44, Appl
37	279	9.8	427	4	US-10-437-963-199493	Sequence 199493,
38	277.5	9.7	668	4	US-10-437-963-134726	Sequence 134726,
39	277.5	9.7	19608	4	US-10-084-846A-8	Sequence 8, Appli
40	274.5	9.6	206	4	US-10-335-977-9083	Sequence 9083, A
41	273.5	9.6	1795	6	US-11-097-143-36210	Sequence 36210, Ap
42	272	9.5	437	4	US-10-437-963-168762	Sequence 168762,
43	268	9.4	1098	5	US-10-450-763-50724	Sequence 50724, A
44	265.5	9.3	555	4	US-10-437-963-203630	Sequence 203630,
45	265	9.3	621	4	US-10-437-963-177697	Sequence 177697,

ALIGNMENTS

RESULT 1
US-10-665-990A-14
; Sequence 14, Application US/10665990A
; Publication No. US20040253222A1
; GENERAL INFORMATION:
; APPLICANT: Apicella, Michael A.
; APPLICANT: Edwards, Jennifer L.
; TITLE OF INVENTION: Vaccine and compositions for the prevention and treatment of Ne
; FILE REFERENCE: 17023-031001
; CURRENT APPLICATION NUMBER: US/10/665,990A
; CURRENT FILING DATE: 2003-09-19
; PRIOR APPLICATION NUMBER: US 10/621,184
; PRIOR FILING DATE: 2003-07-15
; PRIOR APPLICATION NUMBER: US 10/066,551
; PRIOR FILING DATE: 2002-01-31
; PRIOR APPLICATION NUMBER: US 60/344,452
; PRIOR FILING DATE: 2001-10-23
; PRIOR APPLICATION NUMBER: US 60/310,356
; PRIOR FILING DATE: 2001-08-06
; PRIOR APPLICATION NUMBER: US 60/266,070
; PRIOR FILING DATE: 2001-01-31
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14
; LENGTH: 525
; TYPE: PRT
; ORGANISM: Neisseria gonorrhoeae 1291
US-10-665-990A-14

Alignment Scores:
Pred. No.: 6.86e-172 Length: 525
Score: 2689.00 Matches: 520
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0

Query Match:	94.3%	Indels:	0
DB:	5	Gaps:	0
US-10-665-990A-13 (1-1561) x US-10-665-990A-14 (1-525)			
QY	2	AAATACAGGCAATCCGCTCTCAAACTATATCCCGATGAAAACACGAGCGCTCATTTCC	61
DB	6	LysIleGlnAlaMetProSerGluThrIleSerProMetLysThrArgSerLeuIleSer	25
QY	62	CTTTTATGCTCTCTCTGTTTCATGTTCTTCTCATGTTGCCCGCCCACTGGAAGAACGGAGG	121
DB	26	LeuLeuCysLeuLeuLeuCysSerCysSerSerTrpLeuProLeuGluGluArgThr	45
QY	122	GAAGCGCTCATTTCAATTAATCTCCAAACCTGCTCCTCGACAAACATCTGCAATCCGG	181
DB	46	GluSerArgHisPheAsnThrSerLysProValLeuLeuAspAsnIleLeuGlnIleArg	65
QY	182	CACACCCCTATAACACGGGTATCCGACATCTACCTGCTCGACGACCCCGCACGAGCC	241
DB	66	HisThrProHisAsnAsnGlyLeuSerAspIleTyrLeuLeuAspAspProHisGluAla	85
QY	242	CTTGGCGCGCGCGCCCTTATCGAATCTGCGGAACACAGCTCGATTGCAATACTAC	301
DB	86	LeuAlaAlaArgAlaAlaLeuIleGluSerAlaGluHisSerLeuAspLeuGlnTyrTyr	105
QY	302	ATTTGGCGCAACGACATTTCCGGCAGGCTGTTTCAACCTCATGTACCTTCCCGCAGAA	361
DB	106	IleTTrpArgAsnAspIleSerGlyArgLeuLeuPheAsnLeuMetTyrLeuAlaGlu	125
QY	362	CGGGGGTGGCGGTACCGCTGTTGGACGACACACACGCGGGGTGGACCATCTC	421
DB	126	ArgGlyValArgValArgLeuLeuLeuAspAsnAsnThrArgGlyLeuAspAspLeu	145
QY	422	CTGCTCGCCCTCGACAGCCATCCCAATATCGAAGTGGCGCTGTTCAAACCCCTTCGCTTA	481
DB	146	LeuLeuAlaLeuAspSerHisProAsnIleGluValArgLeuPheAsnProPheValLeu	165
QY	482	CCCAATGGCGCGCACTCGGCTACCTGACCGACTTCCCGCGCTCAACCGCGCGATGCAC	541
DB	166	ArgLysTrpArgAlaLeuGlyTyrLeuThrAspPheProArgLeuAsnArgArgMethis	185
QY	542	ACCAATCTTTACCGCGACACCGCGCCACCATCTCGCGGACGCAATATCGCGAC	601
DB	186	AsnLysSerPheThrAlaAspAsnArgAlaThrIleLeuGlyGlyArgAsnIleGlyAsp	205
QY	602	GAATACTTCAAAAGTCGGTCAGACACCGCTTTTCGCGCACTCGACATCTCCGCCACCGC	661
DB	206	GluTyrPheLysValGlyGluAspThrValPheAlaAspLeuAspIleLeuAlaThrGly	225
QY	662	ACGCTCGTCCGGAAGTATCGCACCGACTTTCGACCGCTACTGGGCAAGCCATTCGCGCCAC	721
DB	226	SerValValGlyGluValSerHisAspPheAspArgTyrTrpAlaSerHisSerAlahis	245
QY	722	ACGCGACGCGCATCATCCGACGCGGCACATCGCGCAAGGTCCTCAAGCACTCGGATAC	781
DB	246	AsnAlaThrArgIleIleArgSerGlyAsnIleGlyLysGlyLeuGlnAlaLeuGlyTyr	265
QY	782	AACGACGAAACATCCAGACACCGCTCTCTCGCTACCGCGAAACCGTTCGAAACAGTCGCC	841
DB	266	AsnAspGluThrSerArgHisAlaLeuLeuArgTyrArgGluThrValGluGlnSerPro	285
QY	842	CTCTACCAAAAAATACAGACGGGACGCATCGACTGGCAGAGCGTCCAAACCGCGCTGATC	901
DB	286	LeuTyrGlnLysIleGlnThrGlyArgIleAspTrpGlnSerValGlnThrArgLeuIle	305
QY	902	ACGACACCCCTGCAAAAGGACTCCACCGCGACCGCGCAACCGCGATTCGCGGAGG	961
DB	306	SerAspThrProAlaLysGlyLeuAspArgAspArgGlyProProIleAlaGlyArg	325
QY	962	CTGCAAGACGCGCTCAAAACAGCCCGAAAAAGCGTCTATCTGGTTTACCCCTATTTCGTC	1021
DB	326	LeuGlnAspAlaLeuLysGlnProGluLysSerValTyrLeuValSerProTyrPheVal	345

QY	1022	CCTACAAAATCCGGCACAGACGCACTGGCAAAAACCTGGTGCAAGCGCATAGACGTTACC	1081
DB	346	ProThrLysSerGlyThrAspAlaLeuAlaLysLeuValGlnAspGlyIleAspValThr	365
QY	1082	GTCTGTGCAAACTCGCTACAGCGACCGACGTTGCGCGCTCCATTCCGGCTACGTCAAA	1141
DB	366	ValLeuThrAsnSerLeuGlnAlaThrAspValAlaAlaValHisSerGlyTyrValLys	385
QY	1142	TACCGAAAACCGCTGCTCAAAAGCGGCAATCAAACTCTACGAGCTCAACCCCAACCATGCC	1201
DB	386	TyrArgLysProLeuLeuLysAlaGlyIleLysLeuTyrGluLeuGlnProAsnHisAla	405
QY	1202	GTCCCGCGCCACAAAAGACAAAGCGCTGACCGGACGCTCCGTAAACAGCGCTGCATGCCAAA	1261
DB	406	ValProAlaThrLysAspLysGlyLeuThrGlySerSerValThrSerLeuHisAlaLys	425
QY	1262	ACCTTCATTGTGGACGGCAAAACGCACTTTCATCGGCTCATTCACCACTCGACCCCGTTCC	1321
DB	426	ThrPheIleValAspGlyLysArgIlePheIleGlySerPheAsnLeuAspProArgSer	445
QY	1322	GCACGGCTCAATACCGAAATGGCGTCTGTCATCGAAAGCCCCAAAATCGCAGAACAGATG	1381
DB	446	AlaArgLeuAsnThrGluMetGlyValValIleGluSerProLysIleAlaGluGlnMet	465
QY	1382	GAGCGACCCCTCGCGGATACACACCCCGAATACGCTACCGGTTTACCTCGACAAACAC	1441
DB	466	GluArgThrLeuAlaAspThrThrProGluTyrAlaTyrArgValThrLeuAspLysHis	485
QY	1442	AACGCCCTGCAATGSCAGCATCCGCCACCGCAAAACCTACCCGAAACGAAACCGAAGCC	1501
DB	486	AsnArgLeuGlnTrpHisAspProAlaThrArgLysThrTyrProAsnGluProGluAla	505
QY	1502	AAACTTTGGAAACGCAATCGCGCGCAAAAATCCTATCCCTGCTGCCCATCGAAGTTTATTA	1561
DB	506	LysLeuTrpLysArgIleAlaAlaLysIleLeuSerLeuLeuProIleGluGlyLeuLeu	525
RESULT 2			
US-10-665-990A-16			
; Sequence 16, Application US/10665990A			
; Publication No. US2004025322A1			
; GENERAL INFORMATION:			
; APPLICANT: Apicella, Michael A.			
; APPLICANT: Edwards, Jennifer L.			
; TITLE OF INVENTION: Vaccine and compositions for the prevention and treatment of Neis			
; TITLE OF INVENTION: Infections			
; FILE REFERENCE: 17023-031001			
; CURRENT APPLICATION NUMBER: US/10/665,990A			
; CURRENT FILING DATE: 2003-09-19			
; PRIOR APPLICATION NUMBER: US 10/621,184			
; PRIOR FILING DATE: 2003-07-15			
; PRIOR APPLICATION NUMBER: US 10/066,551			
; PRIOR FILING DATE: 2002-01-31			
; PRIOR APPLICATION NUMBER: US 60/344,452			
; PRIOR FILING DATE: 2001-10-23			
; PRIOR APPLICATION NUMBER: US 60/310,356			
; PRIOR FILING DATE: 2001-08-06			
; PRIOR APPLICATION NUMBER: US 60/266,070			
; PRIOR FILING DATE: 2001-01-31			
; NUMBER OF SEQ ID NOS: 32			
; SOFTWARE: FastSeq for Windows Version 4.0			
; SEQ ID NO 16			
; LENGTH: 525			
; TYPE: PRT			
; ORGANISM: Neisseria gonorrhoeae 1090			
US-10-665-990A-16			
Alignment Scores:			
Pred. No.:	2,36e-171	Length:	525
Score:	2681.00	Matches:	518
Percent Similarity:	99.8%	Conservative:	1
Best Local Similarity:	99.6%	Mismatches:	1
Query Match:	94.0%	Indels:	0
DB:	5	Gaps:	0

US-10-665-990A-13 (1-1561) x US-10-665-990A-16 (1-525)

```
QY 2 AAAATACAGGCAATCGCTCTGAAACTATATATCCCGATGAAACACGAGCTCATTTCC 61
Db 6 LysIleGlnAlaMetProSerGluThrIleSerProMetLysThrArgSerLeuIleSer 25
QY 62 CTTTATGCTCTCTCTCTGTTTCATGTTCTTCTATGTTGCCCCCACTGGAAGACGGAGC 121
Db 26 LeuLeuCysLeuLeuLeuCysSerCysSerSerTrpLeuProLeuGluGluArgThr 45
QY 122 GAAAGCGTCAATTCATCTTCAAACCTGTCCTCTGGACAACATCCTGCAAAATCCGG 181
Db 46 GluSerArgHisPheAsnThrSerLysProValLeuLeuAspAsnIleLeuGlnIleArg 65
QY 182 CACACCCCTCTAAACAGGCTATCCGACATCTACTCTGCTCGAGACCCCGAGAGCC 241
Db 66 HisThrProHisAsnAsnGlyLeuSerAspIleTyrLeuLeuAspAspProHisGluAla 85
QY 242 CTTGCGCGCGCGCGCTTATCGAATCTGCCGAACACAGCTCGATTTCGAATACTAC 301
Db 86 PheAlaAlaArgAlaAlaLeuIleGluSerAlaGluHisSerLeuAspLeuGlnTyrTyr 105
QY 302 ATTTGGCGCAACGACATTTCCGGCAGGCTGCTGTTCAACCTCATGTACCTTGCCTGCA 361
Db 106 IleTrpArgAsnAspIleSerGlyArgLeuLeuPheAsnLeuMetTyrLeuAlaGlu 125
QY 362 CGCGCGTGGCGCTGCTGTTGGACGACACACACGCGCGGGTGGAGCATCTC 421
Db 126 ArgGlyValArgValArgLeuLeuLeuLeuAspAsnAsnThrArgGlyLeuAspLeu 145
QY 422 CTGCTCGCCCTCGACAGCATCCCAATATCGAAGTGGCGCTGTTCAACCCCTTCTGCTCA 481
Db 146 LeuLeuAlaLeuAspSerHisProAsnIleGluValArgLeuPheAsnProPheValLeu 165
QY 482 CGCAAAATGGCGGCACTCGGCTACTGACCGACTTCCCGCCCTCAACCGCGCATGCAC 541
Db 166 ArgLysTrpArgAlaLeuGlyTyrLeuThrAspPheProArgLeuAsnArgArgMetHis 185
QY 542 AACAAATCTTTTACCGCGCAACCGCGCACCATCTACGCGCGGACCAATATCGGCGAC 601
Db 186 AsnLysSerPheThrAlaAspAsnArgAlaThrIleLeuGlyGlyArgAsnIleGlyAsp 205
QY 602 GAATACTTCAAAGTCGGTGAGACACCGTCTTTCGCGCACTGGACATCTCGCCACCGGC 661
Db 206 GluTyrPheLysValGlyGluAspThrValPheAlaAspLeuAspIleLeuAlaThrGly 225
QY 662 AGCGTGTGGCGAAGTATCGACGACTTCGACCGCTACTGGGCAAGCCATTCGCGCCAC 721
Db 226 SerValValGlyGluValSerHisAspPheAspArgTyrTrpAlaSerHisSerAlaHis 245
QY 722 AACGCCACGCGCATCATCGCGCGGCAACATCGCAAGGTCTTCAAGCACTCGGATAC 781
Db 246 AsnAlaThrArgIleIleArgSerGlyAsnIleGlyLysGlyLeuGlnAlaLeuGlyTyr 265
QY 782 AACGACAAACATCAGACACCGCTCTCTGCGCTACCGCGAAACCGTCGAACAGCTCGCCC 841
Db 266 AsnAspGluThrSerArgHisAlaLeuLeuArgTyrArgGluThrValGluGlnSerPro 285
QY 842 CTCCTACCAAAAAATACAGAGCGGACGATCGACTGGCAGAGCGTCCAAAACCCGCTCATC 901
Db 286 LeuTyrGlnLysIleGlnThrGlyArgIleAspTrpGlnSerValGlnThrArgLeuIle 305
QY 902 AGCGACACCCCTGCAAAAGGACTCGACCGCGACCGCGCAACCGCGATTCGCGGAGG 961
Db 306 SerAspSerProAlaLysGlyLeuAspArgAspArgLysProIleAlaGlyArg 325
QY 962 CTGCAAGACGCGCTCAACACAGCCGAAAAAGCGTCTATCTGCTTTTCAACCTATTTCGTC 1021
Db 326 LeuGlnAspAlaLeuLysGlnProGluLysSerValTyrLeuValSerProTyrPheVal 345
QY 1022 CCTACAAAATCCGCGCACAGCGCATCTGGCAAAATCTGTCGAGGACGCGATAGCGTTACC 1081
```

```
Db 346 ProThrLysSerGlyThrAspAlaLeuAlaLysLeuValGlnAspGlyIleAspValThr 365
QY 1082 GTCTGTGACCAACTCGCTACAGGCGACCGACGTTGCGCGCTCCATTCCGGCTACGTCAA 1141
Db 366 ValLeuThrAsnSerLeuGlnAlaThrAspValAlaAlaValHisSerGlyTyrValLys 385
QY 1142 TACCGAAAAACCGCTGCTCAAAAGCGGATCAAACTCTACGAGCTGCAACCCCAACCATGCC 1201
Db 386 TyrArgLysProLeuLeuLysAlaGlyIleLysLeuTyrGluLeuGlnProAsnHisAla 405
QY 1202 GTCCCCGCCCAAAAAGACAAAGGCTGACCGGAGCTCCGTAACACGAGCTCGATGCCAAA 1261
Db 406 ValProAlaThrLysAspLysGlyLeuThrGlySerSerValThrSerLeuHisAlaLys 425
QY 1262 ACCTTCATTGTCGACGCAACGCGATCTTCATTCGGCTCATTAACCTCGACCCCGGTTCC 1321
Db 426 ThrPheIleValAspGlyLysArgIlePheIleGlySerPheAsnLeuAspProArgSer 445
QY 1322 GCACGGCTCAATACCGAAATGGGCGTCTCATCGAAAGCCCAAAATCGCAGAACAGATG 1381
Db 446 AlaArgLeuAsnThrGluMetGlyValValIleGluSerProLysIleAlaGluGlnMet 465
QY 1382 GAGCGCACCTCGCGATACACACCCGAATACGCTTACCGGTTACCTCGACAAACAC 1441
Db 466 GluArgThrLeuAlaAspThrThrProGluTyrAlaTyrArgValThrLeuAspLysHis 485
QY 1442 AACCGCTGCAATGGCAGCATCCCGCCACCGCAAAAACCTACCGAGAACCCGAGAGCC 1501
Db 486 AsnArgLeuGlnTrpHisAspProAlaThrArgLysThrTyrProAsnGluProGluAla 505
QY 1502 AAACCTTTGGAAACGCGATCGCGCAAAATCTATCCCTGCTGCGCATCGAAGGTTTATTA 1561
Db 506 LysLeuTrpLysArgIleAlaAlaLysIleLeuSerLeuLeuProIleGluGlyLeuLeu 525
```

RESULT 3

```
US-10-665-990A-18
; Sequence 18, Application US/10665990A
; Publication No. US20040253222A1
; GENERAL INFORMATION:
; APPLICANT: Apicella, Michael A.
; APPLICANT: Edwards, Jennifer L.
; TITLE OF INVENTION: Vaccine and compositions for the prevention and treatment of Neisseria meningitidis infections
; TITLE OF INVENTION: Infections
; FILE REFERENCE: 17023-031001
; CURRENT APPLICATION NUMBER: US/10/665,990A
; CURRENT FILING DATE: 2003-09-19
; PRIOR APPLICATION NUMBER: US 10/621,184
; PRIOR FILING DATE: 2003-07-15
; PRIOR APPLICATION NUMBER: US 10/066,551
; PRIOR FILING DATE: 2002-01-31
; PRIOR APPLICATION NUMBER: US 60/344,452
; PRIOR FILING DATE: 2001-10-23
; PRIOR APPLICATION NUMBER: US 60/310,356
; PRIOR FILING DATE: 2001-08-06
; PRIOR APPLICATION NUMBER: US 60/266,070
; PRIOR FILING DATE: 2001-01-31
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 18
; LENGTH: 525
; TYPE: PRT
; ORGANISM: Neisseria gonorrhoeae ms11
US-10-665-990A-18

Alignment Scores:
Pred. No.: 4,37e-171 Length: 525
Score: 2677.00 Matches: 517
Percent Similarity: 99.6% Conservative: 1
Best Local Similarity: 99.4% Mismatches: 2
Query Match: 93.9% Indels: 0
DB: 5 Gaps: 0

US-10-665-990A-13 (1-1561) x US-10-665-990A-18 (1-525)
```



```
Db      26  LeuLeuCysLeuLeuLeuCysSerCysSerSerTrpLeuProProLeuGluGluArgThr 45
      |||
Qy      122 GAAAGCCGTCATTTCAATACTTCCAAACCTGTCCTCCCTGGACAAACATCTCGAAATCCGG 181
      |||
Db      46  GluSerArgHisPheAsnThrSerLysProValArgLeuAspAsnIleLeuGlnIleArg 65
      |||
Qy      182 CACACCCCTCATAAACAGGGGTATCCGACATCTACCTGCTCGACGACCCGCCCAAGGCC 241
      |||
Db      66  HisThrProHisThrAsnGlyLeuSerAspIleTyrLeuLeuAsnAspProHisGluAla 85
      |||
Qy      242 CTTCGGCCGCGCGCCCTTATCGAATCTGCGGACACAGCCTCGATTGCAATACTAC 301
      |||
Db      86  PheAlaAlaArgAlaAlaLeuIleGluSerAlaGluHisSerLeuAspLeuGlnTyrTyr 105
      |||
Qy      302 ATTTCGGCGCAACAGCATTTCCGGCAGGCTGCTGTTCAACCTCATGTACCTGCCCGCAGAA 361
      |||
Db      106 IleTrpArgAsnAspIleSerGlyArgLeuLeuPheAsnLeuValTyrLeuAlaAlaGlu 125
      |||
Qy      362 CGCGGGTGGCGGTACGCGCTGCTGTTGGACGACAAACACACGCGGGGTGGACGATCTC 421
      |||
Db      126 ArgGlyValArgValArgLeuLeuLeuAspAsnAsnThrArgGlyLeuAspAspLeu 145
      |||
Qy      422 CTGCTCGCCCTCGACAGCCATCCCAATATCGAAGTGGCGCTGTTCAACCCCTTCGTCCTA 481
      |||
Db      146 LeuLeuAlaLeuAspSerHisProAsnIleGluValArgLeuPheAsnProPheValLeu 165
      |||
Qy      482 CGCAAAATGGCGCGCACTCGGCTACCTGACCGACTTCCCGCGCTCAACCGCGCGATGCAC 541
      |||
Db      166 ArgLysTrpArgAlaLeuGlyTyrLeuThrAspPheProArgLeuAsnArgMetHis 185
      |||
Qy      542 AACAAATCTTTACGCGCGACAAACCGCGCCACCATACTCGGCGGAGCGAATATCGCGAC 601
      |||
Db      186 AsnLysSerPheThrAlaAspAsnArgAlaThrIleLeuGlyGlyArgAsnIleGlyAsp 205
      |||
Qy      602 GAATACTTCAAGTGGTGAGACACCGTTTCGCGCACTCGACATCTCGCATCTCGCCACCGC 661
      |||
Db      206 GluTyrPheLysValGlyGluAspThrValPheAlaAspLeuAspIleLeuAlaThrGly 225
      |||
Qy      662 AGCGTGTGCGGGAAGTATCGCACGCACTTCGACCGCTACTGGCGAAGCCATTTCCGCCAC 721
      |||
Db      226 SerValValGlyGluValSerHisAspPheAspArgTyrTrpAlaSerHisSerAlaHis 245
      |||
Qy      722 AACGCCACGCGCATCATCGCGAGCGGCAACATCGGCAAGGGTCTTCAAGCACTCGGATAC 781
      |||
Db      246 AsnAlaThrArgIleIleArgSerGlyAsnIleGlyLysGlyLeuGlnAlaLeuGlyTyr 265
      |||
Qy      782 AACGAGAAACATCCAGACACGCGCTCTCGCGCTACCGGAAACCGTCGAAACAGTCGCC 841
      |||
Db      266 AsnAspGluThrSerArgHisAlaLeuLeuArgTyrArgGluThrValGlnSerPro 285
      |||
Qy      842 CTCTACAAAAAATACAGACGGGACGCATCGACTGCGAGAGCGTCCAAACCCGCTGATC 901
      |||
Db      286 LeuTyrGlnLysIleGlnThrGlyArgIleAspTrpGlnSerValGlnThrArgLeuIle 305
      |||
Qy      902 ACGCACACCCCTGCAAAAGGACTCGACCGCGACCGCGCAACCGCGGATTCGCGGAGG 961
      |||
Db      306 SerAspAspProAlaLysGlyLeuAspArgAspArgLysProProIleAlaGlyArg 325
      |||
Qy      962 CTGCAAGACGCGTCAAAACAGCCCGAAAAAGCGTCTATCTGGTTTACCCCTATTTCGTC 1021
      |||
Db      326 LeuGlnAspAlaLeuLysGlnProGluLysSerValTyrLeuValSerProTyrPheVal 345
      |||
Qy      1022 CCTACAAAATCCGGCACACAGCACTGGCAAAACTGGTCGACGAGCGGCATAGACGTTACC 1081
      |||
Db      346 ProThrLysSerGlyThrAspAlaLeuAlaLysLeuValGlnAspGlyIleAspValThr 365
      |||
Qy      1082 GTCTTGACCAACTCGCTACAGGCGACCGAGTTCGCGCGCTCCATTCGGCTACGTCAAA 1141
      |||
Db      366 ValLeuThrAsnSerLeuGlnAlaThrAspValAlaAlaValHisSerGlyTyrValLys 385
      |||
Qy      1142 TACCGAAAAACCGCTGCTCAAAGCCGGCATCAAACTCTACGAGCTGCAACCCCAACATGCC 1201
      |||
```

```
Db      386 TyrArgLysProLeuLeuLysAlaGlyIleLysLeuTyrGluLeuGlnProAsnHisAla 405
      |||
Qy      1202 GTCCCCGCCCAAAAGACAAAAGCGCTGACCGGACGCTCCGTAACAGAGCTGCATGCCAAA 1261
      |||
Db      406 ValProAlaThrLysAspLysGlyLeuThrGlySerSerValThrSerLeuHisAlaLys 425
      |||
Qy      1262 ACCTTCATTCGGAGCGCAAAAGCATCTTCATCGGCTCATTCACCTCGACCCCGTTC 1321
      |||
Db      426 ThrPheIleValAspGlyLysArgIlePheIleGlySerPheAsnLeuAspProArgSer 445
      |||
Qy      1322 GCACGGCTCAATACCGAAATGGCGTCGTCATCGAAAGCCCCAAAATCGCAGAACAGATG 1381
      |||
Db      446 AlaArgLeuAsnThrGluMetGlyValValIleGluSerProLysIleAlaGluGlnMet 465
      |||
Qy      1382 GAGCGCACCTCGCGCATACCCACACCCCGAATACGCTACCGGTTTACCCTCGACAAACAC 1441
      |||
Db      466 GluArgThrLeuAlaAspThrSerProGluTyrAlaTyrArgValThrLeuAspArgHis 485
      |||
Qy      1442 AACCGCTGCAATGGCAGCATCCCGCCACCGAAACCTACCGAAACCGAACCCGAGCC 1501
      |||
Db      486 AsnArgLeuGlnTrpHisAspProAlaThrArgLysThrTyrProAsnGluProGluAla 505
      |||
Qy      1502 AAACTTTCGAAACGCGATCCCGCAAAAATCCTATCCCTGCTGCCCATCGAAGGTTTATTA 1561
      |||
Db      506 LysLeuTrpLysArgIleAlaAlaLysIleLeuSerLeuLeuProIleGluSerLeuLeu 525
      |||

RESULT 6
US-10-665-990A-20
; Sequence 20, Application US/10665990A
; Publication No. US20040253222A1
; GENERAL INFORMATION:
; APPLICANT: Apicella, Michael A.
; APPLICANT: Edwards, Jennifer L.
; TITLE OF INVENTION: Vaccine and compositions for the prevention and treatment of Neis
; TITLE OF INVENTION: Infections
; FILE REFERENCE: 17023-031001
; CURRENT APPLICATION NUMBER: US/10/665,990A
; CURRENT FILING DATE: 2003-09-19
; PRIOR APPLICATION NUMBER: US 10/621,184
; PRIOR FILING DATE: 2003-07-15
; PRIOR APPLICATION NUMBER: US 10/066,551
; PRIOR FILING DATE: 2002-01-31
; PRIOR APPLICATION NUMBER: US 60/344,452
; PRIOR FILING DATE: 2001-10-23
; PRIOR APPLICATION NUMBER: US 60/310,356
; PRIOR FILING DATE: 2001-08-06
; PRIOR APPLICATION NUMBER: US 60/266,070
; PRIOR FILING DATE: 2001-01-31
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 20
; LENGTH: 525
; TYPE: PRT
; ORGANISM: Neisseria meningitidis B
US-10-665-990A-20

Alignment Scores:
Pred. No.: 6,21e-168 Length: 525
Score: 2630.00 Matches: 509
Percent Similarity: 98.7% Conservative: 4
Best Local Similarity: 97.9% Mismatches: 7
Query Match: 92.2% Indels: 0
DB: 5 Gaps: 0

US-10-665-990A-13 (1-1561) x US-10-665-990A-20 (1-525)
Qy      2  AAAATACAGGCAATCCGTCGTGAAACTATATCCCCGATGAAACACGAGCGCTCATTTCC 61
      |||
Db      6  LysThrGlnAlaMetProSerGluThrIleSerLeuMetLysThrArgSerLeuLeuSer 25
      |||
Qy      62  CTTTATTCGCTCTCTCTGTTTCATGTTCATCGTTCGCCCCCACTGGAAAGACGAGCG 121
      |||
Db      26  LeuLeuCysLeuLeuLeuCysSerCysSerSerTrpLeuProProLeuGluGluArgThr 45
      |||
```

QY 122 GAAAGCGTCAATTTCAATCTTCCAAACCTGTCTCTCCGGACCAACATCTCTGCAAAATCCGG 181
DB 46 GluSerArgHisPheAsnThrSerLysProValArgLeuAspAsnIleuGlnIleArg 65
QY 182 CACACCCCTCAACAAACGGGTATCCGACATCTACCTGTCTGACGACCCCAACAGCC 241
DB 66 HisThrProHisThrAsnGlyLeuSerAspIleTyrLeuLeuAsnAspProHisGluAla 85
QY 242 CTTGGCCGCGCGCCCTTATCGAATCTGCGAACACAGCCTCGATTTCGAATCTAC 301
DB 86 PheAlaAlaArgAlaLeuIleGlnSerAlaGluHisSerLeuAspLeuGlnTyrTyr 105
QY 302 ATTTGGCCCAACGACATTTCCGGCAGGTGTGTTCAAACCTCATCTACCTTTCGCCGACAA 361
DB 106 IleTyrArgAsnAspIleSerGlyArgLeuLeuPheAsnLeuValTyrLeuAlaGlu 125
QY 362 CGCGCGTGGCGGTACGCTGTGTGGACGACAAACAAACGCGGGTTCGACATCTC 421
DB 126 ArgGlyValArgValArgLeuLeuAspAspAsnAsnThrArgGlyLeuAspLeu 145
QY 422 CTGCTCGCCCTCGACAGCATCCCAATATCGAAGTGGCCTGTTCACACCCCTTCGTCCTA 481
DB 146 LeuLeuAlaLeuAspSerHisProAsnIleGluValArgLeuPheAsnProPheValLeu 165
QY 482 CGCAATGGCGGCACTCGGCTACCTGACCGACTTCCCGCCGCTCAACCGCGCATGCAC 541
DB 166 ArgLysTyrArgAlaLeuGlyTyrLeuThrAspPheProArgLeuAsnArgMethHis 185
QY 542 AACAAATCTTTACCGCGACACACCGCGCCACCTATCTCGCGGACGCAATATCGGAC 601
DB 186 AsnLysSerPheThrAlaAspAsnArgAlaThrIleLeuGlyGlyArgAsnIleGlyAsp 205
QY 602 GAATACTTCAAGTCGGGAGACACCGCTTTCGCGACCTCGACATCTCGCCACCGCC 661
DB 206 GluTyrPheLysValGlyGluAspThrValPheAlaAspLeuAspIleLeuAlaThrGly 225
QY 662 AGCGTCTGCGGGAAGTATCGCACGACTTCGACCGCTACTGGGCAAGCATTCGCGCCAC 721
DB 226 SerValValGlyGluValSerHisAspPheAspArgTyrTrpAlaSerHisSerAlaHis 245
QY 722 AACGCGACGCGATCATCCGAGCGGCAACATCGGCAAGGTCTTCAAGCACTCGGATAC 781
DB 246 AsnAlaThrArgIleIleArgSerGlyAsnIleGlyLysGlyLeuGlnAlaLeuGlyTyr 265
QY 782 AACGAGAAACATCCAGACGCGCTCTCGCTACCGGAAACCGTCAACACATCGCC 841
DB 266 AsnAspGluThrSerArgHisAlaLeuLeuArgTyrArgGluThrValGluGlnSerPro 285
QY 842 CTCTACCAAAAAATACAGACGGGACGCAATCGACTGGCAGAGCGTCCAAACCCGCTGATC 901
DB 286 LeuTyrGlnLysIleGlnThrGlyArgIleAspTrpGlnSerValGlnThrArgLeuIle 305
QY 902 AGCAGACCCCTCGAAAGAGTCTGACGCGACCGCCGCAACCCGCGATTCCGGGAGG 961
DB 306 SerAspAspProAlaLysGlyLeuAspArgAspArgLysProIleAlaGlyArg 325
QY 962 CTGCAAGACGCTCAACACGCGCCGAAAAAGCGCTCTATCTGGTTTACCCCTATTTCGTC 1021
DB 326 LeuGlnAspAlaLeuLysGlnProGluLysSerValTyrLeuValSerProTyrPheVal 345
QY 1022 CCTCAAAATCCGGCACAGACGCACTGGCAAACTGGTGCAGGACGGCATAGAGCTTACC 1081
DB 346 ProThrLysSerGlyThrAspAlaLeuAlaLysLeuValGlnAspGlyIleAspValThr 365
QY 1082 GTCCTGACCAACTCGCTACAGGACCGACGCTTCGCCCGCTCCATTCCGGCTACGTCAAA 1141
DB 366 ValLeuThrAsnSerLeuGlnAlaThrAspValAlaAlaHisSerGlyTyrValLys 385
QY 1142 TACGAAACCGCTCTCAAGCGGATCAAACTCTACGAGCTGCAACCAACCATGCC 1201
DB 386 TyrArgLysProLeuLeuLysAlaGlyIleLysLeuTyrGluLeuGlnProAsnHisAla 405

QY 1202 GTCCCGCCCAAAAGAGCGCTGACCGGAGCTCCGTAAACAGCGCTCATGCCAAA 1261
DB 406 ValProAlaThrLysAspLysGlyLeuThrGlySerSerValThrSerLeuHisAlaLys 425
QY 1262 ACCTTCATTGTGGACGCAAAACGACATCTTCATCGGCTCAITCAACCTCGACCCCGTTCC 1321
DB 426 ThrPheIleValAspGlyLysArgIlePheIleGlySerPheAsnLeuAspProArgSer 445
QY 1322 GCAGGCTCAATACGGAATGGCGTGTCTATCGAAGCCCAAAATCGAGACAGATG 1381
DB 446 AlaArgLeuAsnThrGluMetGlyValIleGluSerProLysIleAlaGluGlnMet 465
QY 1382 GAGCGCACCTCGCGCATACACACCGAATACCGCTACCGCTTACCTCGACAAAACAC 1441
DB 466 GluArgThrLeuAlaAspThrSerProGluTyrAlaTyrArgValThrLeuAspArgHis 485
QY 1442 AACCGCTGCAATGGCAGATCCCGCCACCGCAAAACCTACCCGAAACGCAAGCC 1501
DB 486 AsnArgLeuGlnTyrHisAspProAlaThrArgLysThrTyrProAsnGluProGluAla 505
QY 1502 AAACCTTTGGAAACGATCGCGCGCAAAATCTATCCCTGCTGCCCATCGAAGTTTATTA 1561
DB 506 LysLeuTrpLysArgIleAlaAlaLysIleLeuSerLeuLeuProIleGluSerLeuLeu 525

RESULT 7
US-10-335-977-9085
; Sequence 9085, Application US/10335977
; Publication No. US20040052799A1
; GENERAL INFORMATION:
; APPLICANT: DOUGLAS SMITH et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES
; RELATING TO HELICOBACTER PYLORI FOR
; DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 10031
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 28 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: Windows NT 4.0
; SOFTWARE: UNIX
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/335,977
; FILING DATE: 30-Dec-2002
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/993,002
; FILING DATE: 17-DEC-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Mandragouras, Amy E.
; REGISTRATION NUMBER: 36,207
; REFERENCE/DOCKET NUMBER: CTN-018
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)742-4214
; INFORMATION FOR SEQ ID NO: 9085:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 502 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: Helicobacter pylori
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (B) LOCATION 1...502
; SEQUENCE DESCRIPTION: SEQ ID NO: 9085:
US-10-335-977-9085

; INFORMATION FOR SEQ ID NO: 9086:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 502 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; HYPOTHETICAL: YES
 ; ORIGINAL SOURCE:
 ; ORGANISM: Helicobacter pylori
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; LOCATION: (B) LOCATION 1...502
 ; SEQUENCE DESCRIPTION: SEQ ID NO: 9086:
 US-10-335-977-9086

Alignment Scores:

Pred. No.: 6, 25e-34 Length: 502
 Score: 632.00 Matches: 161
 Percent Similarity: 51.7% Conservative: 93
 Best Local Similarity: 32.8% Mismatches: 181
 Query Match: 22.2% Indels: 56
 DB: 4 Gaps: 17

US-10-665-990A-13 (1-1561) x US-10-335-977-9086 (1-502)

QY 176 ATCCGGCACACCCCTCATACAAACGGGCTATCCGACATCTAC----- 217
 DB 30 lIeSerTyAspProTyThrThrThrIleGlySerLeuTyAlaLysAsnLeuLysGlu 49
 QY 218 -----CTGCTCAGACACCCCAAGAGCCCTTCCGCCCGC 253
 DB 50 AsnProLysHisSerAlaAlaIleLeuLeuGluAspGlyPheAspAlaLeuHisArg 69
 QY 254 GCCGCCCTTATCGAATCGCGAACACACAGCTTCGATTTGCAATACTACATTTGGCGCAAC 313
 DB 70 ValGlyLeuIleArgMetSerGlnLysSerIleAspMetGlnThrTyrlleTyLysAsn 89
 QY 314 GACATTTCCGGCAGGCTGCTTTCAACCTCATCTTGGCGGAGAACGGGGTGGC 373
 DB 90 AspLeuSerSerGlnValIleAlaLysGluLeuLeuAsnAlaAlaAsnArgGlyValLys 109
 QY 374 GTACGCTGCTGTTGGACGACAAACACACGCGCGGTGGAC-----GATCTC 421
 DB 110 ValArgIleLeuLeuAspAsn-----GlyLeuAspSerAspPheSerAspIle 126
 QY 422 CTGCTCGCCCTCGACAGCATCCCAATATCGAAGTGGCGCTTCAACCCCTTCGTCTTA 481
 DB 127 MetLeu---LeuAsnPheHisLysAsnIleGluValLysIlePheAsnProTyrlle 145
 QY 482 CGCAATGGCGCGCACTCGGTAC-----CTGACCGACTTCCCGCGCTCAACCGC 532
 DB 146 ArgAsn---LysGlyLeuArgTyThrPheGluMetLeuAlaAspTyThrGluArgIleLysLys 164
 QY 533 CGCATGCACAAATCTTATCCCGCGACAAACCGCGCCACCATCTCGCGGAGCGCAAT 592
 DB 165 ArgMetHisAsnLysLeuPheIleValAspAsnPheAlaValIleIleGlyArgAsn 184
 QY 593 ATCCGCGACGAATCTCAAGTCGTCGAGACACCGTTTTCGCGGACCTGGACATCTC 652
 DB 185 IleGlyAspAsnTyThrPheAspAsnAspLeuAspThrAsnPheLeuAspLeuAspAlaLeu 204
 QY 653 GCCACCGCGACGCTCGTCGGCGAAGTATCGACGACTTCGACGCTTCTGGCAAGCAT 712
 DB 205 PhePheGlyGlyValAlaSerLysAlaLysGluSerPheGluAsnTyThrArgPheHis 224
 QY 713 TCCGCCCAACAGCCGCGCATCATCCGACG-----GGC 748
 DB 225 ArgSerIleProValSer---LeuLeuArgThrHisLysArgLeuLysAsnAsnValLys 243
 QY 749 AACATCGCGAGGCTCTCAAGCACTCGGATACACGACGAAACATCCAGACACGCGCTC 808
 DB 244 GluIleAlaLysLeuHisGluLysIleProIleSerAlaGluAspAlaAsnGluPheGlu 263

QY 809 CTGGCTACCGCGAAACCGTCCAAACAGTCGCGCCCTCTACCAA---AAAATACAGACGGCA 865
 DB 264 LysLysValAsnAspPheIleGluArgPheGlnLysTyThrGlnTyThrProIleTyThrGly 283
 QY 866 CGCATCGACTGGCAGAGCTCCAAACCCGCTGTATCAGCAGACCCCTGCAAAAGGACTC 925
 DB 284 -----AsnAlaIlePheLeuAlaAspLeuProAlaLys---Ile 295
 QY 926 GACCGCGACCGCGCAACCGCATTCGCGGAGGCTGCAAGACGCGCTCAAAACAGGCC 985
 DB 296 AspThrProLeuTyThrSer---ProIleLysIleAlaPheGluLysAlaLeuLysAsnAla 314
 QY 986 GAAAAAGCGTCTATCTGGTTTACCCCTATTCCTCCTACAAATCCGCGACAGACGCA 1045
 DB 315 LysAspSerValPheIleAlaSerSerTyThrPheIleProGlyLysLysIleMetLysIle 334
 QY 1046 CTGGCAAACTGGTGCAGGAGCGATAGACGTTACCGTCTCTGACCAACTCCTCAGCGC 1105
 DB 335 PheLysAsnGlnIleSerLysGlyIleGluLeuAsnIleLeuThrAsnSerLeuSerSer 354
 QY 1106 ACCGACGTTGCGCGCGCTCCATTCGCGCTACGTCAATACCGAAACCGCTGCTCAAGCC 1165
 DB 355 ThrAspAlaIleValValTyThrGlyAlaTrpGluArgTyArgAsnLysLeuValArgMet 374
 QY 1166 GGCAATCAAACTTACGAGCTGCAACCAACCATCCGCTCCCGCCACAAAGACAAAGCC 1225
 DB 375 GlyAlaAsnValTyGluIleArgAsnAspPhePheAsnArgGlnIleLysGlyArg--- 393
 QY 1226 CTGACCGGAGCTCGTAACAGCAGCTGCATGCCAAAACCTTTCATTGTGGACGCGCAACGC 1285
 DB 394 -----PheSerThrLysHisSerLeuHisGlyLysThrIleValPheAspAlaLeu 411
 QY 1286 ATCTTCATCGCTCATTCACCTCGACCCCGTTCGCGACGCTCAATACCGAAATGGCG 1345
 DB 412 ThrLeuLeuGlySerPheAsnIleAspProArgSerAlaTyIleAsnThrGluSerAla 431
 QY 1346 GTCTCATCAAAAGCCCAAAATCGCAGAACAGATGGAGGCGACCCCTCGCGATACCA 1405
 DB 432 ValLeuPheAsnProSerPheAlaLysArgValArgLeuSerLeuLysAspHisAla 451
 QY 1406 CCGCAATACGCTACCGCGTTACCTCGACAAACACACCGCTGCAATACGACGATCCC 1465
 DB 452 ---GlnGlnSerThrHisLeuValLeuTyArgHis---ArgValIleTrpGlu----- 467
 QY 1466 GCCACCGCAAAACCTAC-----CCGAAACGAAACCGAAGCCAACTTTGGAAA 1513
 DB 468 AlaThrGluGluGlyIleLeuIleHisGluLysAsnSerProAspThrSerPhePheLeu 487
 QY 1514 CGCATCGCGCAAAATCCTATCCCTGCTGCC 1546
 DB 488 ArgLeuIleLysGluTrpSerLysValLeuPro 498

RESULT 9

US-09-881-752A-356
 ; Sequence 356, Application US/09881752A
 ; Patent No. US20020115078A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Kleanthous, Harold
 ; APPLICANT: Al-Garawi, Amal
 ; APPLICANT: Miller, Charles
 ; APPLICANT: Tomb, Jean-Francois
 ; APPLICANT: Oomen, Raymond P.
 ; TITLE OF INVENTION: Identification of polynucleotides
 ; TITLE OF INVENTION: Encoding No. US20020115078A1 Helicobacter Polypeptides in the
 ; FILE REFERENCE: 06132/041002
 ; CURRENT APPLICATION NUMBER: US/09/881,752A
 ; PRIOR FILING DATE: 2001-06-15
 ; PRIOR APPLICATION NUMBER: US 08/833,457
 ; NUMBER OF SEQ ID NOS: 370
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 356

```
; LENGTH: 428
; TYPE: PRT
; ORGANISM: Helicobacter pylori
US-09-881-752A-356

Alignment Scores:
Pred. No.:      3,996-31      Length:      428
Score:          590.00      Matches:      145
Percent Similarity: 50.1%      Conservative: 80
Best Local Similarity: 32.3%      Mismatches: 146
Query Match:      20.7%      Indels:      78
DB:               3         Gaps:      14

US-10-665-990A-13 (1-1561) x US-09-881-752A-356 (1-428)

QY 272 GCGCAACACAGCCTCGATTGCAATACATACATTTGGCGCAACAGACATTTCCGGCAGGCTG 331
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2 SerGlnLysSerIleAspMetGlnThrTyrIleTyrLysAsnAspSerSerGlnVal 21

QY 332 CTGTTCAACCTCATGTTACCTTCCGCGAGAACCGCGGTCGCGTACGCTGCTGTGGAC 391
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 22 IleAlaLysGluLeuLeuAsnAlaAlaAsnArgGlyValLysValArgIleLeuLeuAsp 41

QY 392 GACAAACACACGCGGGGTGCAC-----GATCTCTGCTCGCCCTCGACAGC 439
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 42 AspAsn-----GlyLeuAspSerAspPheSerAspIleMetLeu---LeuAsnPhe 57

QY 440 CATCCCAATATCGAAGTGGCTGTTCACCCCTTCCTCTACGCAAAATGGCGCGCACTC 499
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 58 HisLysAsnIleGluValLysIlePheAsnProTyrTyrIleArgAsn---LysGlyLeu 76

QY 500 GGCTAC-----CTGACCGACTTCCCGCCCTCAACCGCGGATGACCAACAAATCC 550
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 77 ArgTyrPheGluMetLeuAlaAspTyrGluArgIleLysLysArgMetHisAsnLysLeu 96

QY 551 TTTACCGCGCAACCGCCACCATCTCGCGGAGCGCAATATCGCGCAGCAATACTTC 610
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 97 PheIleValAspAsnPheAlaValIleIleGlyArgAsnIleGlyAspAsnTyrPhe 116

QY 611 AAGTCGGTGAGCACCACCGTTTCGCGACCTGGACATCTCGCCACCGCGGACGCTGCTC 670
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 117 AspAsnAspLeuAspThrAsnPheLeuAspLeuAspAlaLeuPheGlyValAla 136

QY 671 GCGAAGTATCCAGCATCTCGACGCTACTCGGCAAGCATCTCGCCCAACACCGCAGC 730
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 137 SerLysAlaLysGluSerPheGluArgTyrTrpArgPheHisArgSerIleProValSer 156

QY 731 CGCATCATCGCGAGCGCAACATCGCAAGGCTCTCAAGCATCGGATACAACGACGAA 790
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 157 ----LeuLeuArgThr----- 160

QY 791 ACATCCAGACACGCGCTCCTCGCTACCGCGAACCCTGGAACAGTCGCGCCCTTACCAA 850
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 161 -----HisLysArgLeuLys---AsnAsnAlaLysGluIleAlaLysLeuHisGlu 176

QY 851 AAAATACACAGCGGACGATCACTGGCAGAGCGTCCAAACCGCGCTG----- 898
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 177 LysIleProIleSerAlaGluAspLysAsnGlnPheGluLysLysValAsnAspPheIle 196

QY 899 ----- 907
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 197 AspArgPheGlnLysTyrGlnTyrProIleTyrTyrGlyAsnAlaIlePheLeuAlaAsp 216

QY 908 ACCCTGCAAAAGACTCGACCGCGACCGCCGCAACCGCGGATGCGGGAGGCTGCAAA 967
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 217 SerProLysLys---IleAspThrProLeuTyrSer---ProIleLysIleAlaPheGlu 234

QY 968 GACGGCTCAACACAGCCCAAAAGCGTCTATCTGGTTTCACCTATTTCTCCCTACA 1027
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 235 LysAlaLeuLysAsnAlaLysAspSerValPheIleAlaSerSerTyrPheIleProGly 254

QY 1028 AAATCCGGCAGACGACGACTGGCAAAACTGGTGCAGGACGCGCATAGACGTTACCGTCTG 1087
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

```
Db 255 LysLysMetMetLysIlePheLysAsnGlnLysSerLysGlyIleGluLeuAsnIleLeu 274
QY 1088 ACCAACTCGTACAGCGACGCGATTCGCCCGCTTCATTCGGCTACGTCAAATACCGA 1147
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 275 ThrAsnSerLeuSerSerThrAspAlaIleValValTyrGlyAlaTyrGluArgTyrArg 294

QY 1148 AAACCGCTGCTCAAAAGCGGATCAAACTCTACGAGCTGCAACCAACCATGCGCTCCC 1207
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 295 AsnGlnLeuValArgMetGlyAlaAsnValTyrGluIleArgAsnAspPhePheAsnArg 314

QY 1208 GCCACAAAGACAAAGCGCTGACCGGACGCTCCGTAACACGAGCTCATGCCAAAACCTTC 1267
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 315 GlnIleLysGlyArg-----PheSerThrLysHisSerLeuHisGlyLysThrIle 331

QY 1268 ATTGTGGACGGCAACGCGCATCTTCATCGGCTCATTCACCTCGACCCCGCTTCGACGG 1327
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 332 ValPheAspAspAsnLeuThrLeuLeuGlySerPheAsnIleAspProArgSerAlaTyr 351

QY 1328 CTCAATACCGAAATGGGCGTGTCTATCGAAAGCCCCAAAATCGCAGACAGATGGAGCGC 1387
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 352 IleAsnThrGluSerAlaValLeuPheAspAsnProSerPheAlaLysArgVal----- 369

QY 1388 ACCCTCGCGCATACACACCCCGAATACGCTACCGCTTACCTCGACAAACACAAACGCG 1447
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 370 -----ArgLeuSerLeuLysAspHisAlaGln 378

QY 1448 CTGCAATGCGACGATCCGCGCACCGCAAAAACCTACCCGAAACGAAACCGAGCCAACTT 1507
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 379 GlnSerTyrHisLeuValValTyrArg-----HisArgValIle 391

QY 1508 TGGAAACGATCGCGCGCAAAAATCCTA 1534
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 392 TrpGluAlaValGluGluGlyIleLeu 400

RESULT 10
US-10-335-977-9084
; Sequence 9084, Application US/10335977
; Publication No. US20040052799A1
; GENERAL INFORMATION:
; APPLICANT: DOUGLAS SMITH et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES
; RELATING TO HELICOBACTER PYLORI FOR
; DIAGNOSTICS AND THERAPEUTICS
;
; NUMBER OF SEQUENCES: 10031
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 28 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: Windows NT 4.0
; SOFTWARE: UNIX
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/335,977
; FILING DATE: 30-Dec-2002
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/993,002
; FILING DATE: 17-DEC-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Mandragouras, Amy E.
; REGISTRATION NUMBER: 36,207
; REFERENCE/DOCKET NUMBER: GTN-018
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)742-4214
; INFORMATION FOR SEQ ID NO: 9084:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 224 amino acids
; TYPE: amino acid
```


Qy	461	CTGTTCAACCCGTCCTAGGCAAAATGGCGGCACTCGGCTACCTGACCGACTTCGCC	520
Db	1423	---SerProThrThrThrThrThrThrProProProThrThrThrProSerProPr	1441
Qy	521	CGCCTCAACCGCGCATGCACAAACAAATCCCTTTACCGCGCAACACCGCCACCATACTC	580
Db	1441	oileThrThr-----ThrThrProLeuProThrThrThrProSerPro-----	1456
Qy	581	GGCGGACGCAATATCGGCGGACGAATACTTCAAATCGGTGAGGACACCGTTTCGCCGAC	640
Db	1457	-----ProIleSerThrThrThrProProProThrThrThrThrProProProTh	1474
Qy	641	CTGGACATCTCGCCACCGGCGAGCGTCGTCGGCGAAGTATCGCAGGACTTCGACC-----	695
Db	1474	rThrThrProProProThrThrThrProSerProProThrThrThrThrThrThrPr	1494
Qy	696	-----GCTACTGGGAAGCGATTTCGGCCCAACACCGCCACGCGCATCATCCGCGCGGC	748
Db	1494	oProProThrThrThrProSerProProMetThrThrProIleThrProProAlaSerTh	1514
Qy	749	AACATCGGCAAGGCTTCAAGCACTCGGATACACGACGACGAAACATCCAGACGCGCTC	808
Db	1514	rThrThrLeu-----ProProThrThrThrProSerProProThrThrTh	1529
Qy	809	CTGCGCTACCGGAAACCGTCGAAC-----AGTCGGCCCTCTACCAAAATAACACAGC	862
Db	1529	rThrThrThrProProProThrThrThrProSerProProThrThr-----	1544
Qy	863	GGACGCATCGACTGGCAGAGCGTCCAAACCCGCGCTGATCGACGACACCCCTGTCAAAAGGA	922
Db	1545	-----ThrProIleThrProPr	1550
Qy	923	CTGACCGCGACCGCGGAAACCGCGATTGTCGGGAGGTGCAAGACGCGCTCAACACAG	982
Db	1550	oSerThrThrThr-----LeuPro-----ProThrThrThrProSe	1563
Qy	983	CCCGAAAAAGCGTCTATCTGGTTTACCCCTATTTCGTCCTACAAATCGGCACAGAC	1042
Db	1563	rProProProThrThrThrThrThrProProProThrThrThrProSerProProThrTh	1583
Qy	1043	GCACTGCCAAAACTGGTGCAGGACGATAGACGTTACCGTCTCTGACCA-----AC	1093
Db	1583	r-----ThrThrProSerProProThrIleThrTh	1593
Qy	1094	TCGCTACAGGGACGAGTTGCGCGCGTCGATTCGGCTACGTCAAT-----	1142
Db	1593	rThrThrProProProThrThrThrProSerProProThrThrThrThrThrThrProPr	1613
Qy	1143	-----ACGAAAAACCGTGTCTCAAGCCGCGATCAAACTC	1177
Db	1613	oProThrThrThrProSerProProThrThrThrProIleThrProProThrSerThrTh	1633
Qy	1178	TACGAGCTGCAACCCAACTATCGCGTCCCGGCCCAACAAAGACACGCGTTCGCGGACG	1237
Db	1633	rThrLeuProProThrThrThrProSerProProProThr-----ThrThrTh	1650
Qy	1238	TCGGTAACGACCTGCATGCGAAACCTTCATTTGGACGCGCAACGCATCTTCATCGCGC	1297
Db	1650	rProProProThrThrThrProSerProProThr---ThrThrThrProSerProPrI	1669
Qy	1298	TCATTCAACCTCGACCCCGTTCCGACGGCTCAATACCGAAATGGCGTCGTCATCGAA	1357
Db	1669	eThrThrThrThrProProProThrThrThrProSerProIleThrThrThrPr	1689
Qy	1358	AGCCCCAAAATCGCAGAACAGATGGAGCGCACCCCTCGCCGATACCAACCCGGAATACGCC	1417
Db	1689	oSerProProThr---ThrThrMetThrThrProSerProThrThrThrProSerPr	1708
Qy	1418	TACCGGTTACCTCGCAACAAACACCGCGCTGCAATGGACGATCCCGCCACCCGAA	1477
Db	1708	oile-----ThrThrThrThrProSerSerThrThrThrProSerProPr	1725

```

Qy 1478 ACCTACCGGACGACCAACCGAAGCCAAACCTTTGGAAACGCATCGCCGCAAAAATCCTTATCC 1537
Db 1725 oThrMetThrThrProSerProThrThrThrProSerProThrThrThrMetTh 1745
Qy 1538 CTGCTGCCCCA 1547
Db 1745 rThrLeuPro 1748

RESULT 13
US-10-025-380-1068
; Sequence 1068, Application US/10025380
; Publication No. US20020182191A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Lodes, Michael J.
; APPLICANT: Secrist, Heather
; APPLICANT: Benson, Darin R.
; APPLICANT: Meagher, Madeleine Joy
; APPLICANT: Stolk, John A.
; APPLICANT: Wang, Tongtong
; APPLICANT: Jiang, Yugu
; APPLICANT: Smith, Carole L.
; APPLICANT: King, Gordon E.
; APPLICANT: Wang, Aijun
; APPLICANT: Clapper, Jonathan D.
; APPLICANT: Skeiky, Yasir A. W.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedvick Thomas S.
; APPLICANT: Carter, Darrick
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND DIAGNOSIS
; TITLE OF INVENTION: OF COLON CANCER AND METHODS FOR THEIR USE
; FILE REFERENCE: 210121.471C14
; CURRENT APPLICATION NUMBER: US/10/025,380
; CURRENT FILING DATE: 2001-12-19
; NUMBER OF SEQ ID NOS: 1129
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1068
; LENGTH: 5179
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-025-380-1068

Alignment Scores:
Pred. No.: 6,05e-14 Length: 5179
Score: 336.50 Matches: 141
Percent Similarity: 33.6% Conservative: 55
Best Local Similarity: 24.1% Mismatches: 198
Query Match: 11.8% Indels: 190
DB: Gaps: 22

US-10-025-380-1068 (1-1561) x US-10-025-380-1068 (1-5179)
Qy 119 ACGGAAGCGGTCAATTTCAAT-----ACTTCAAACCTGCTCTCTGGACAAC 166
Db 1248 ThrValGlulysHisPheAsnIleCysSerIleThrThrArgPro-SerThrLeuThrTh 1267
Qy 167 ATCTCGCAATCGGCACACCCCTCATACAACGGGCTATCCGACATCTACCTGCTCGAC 226
Db 1267 rPheThrThrIleThrLeuLeuProThrThrProThrSerPheThrThrThrThrTh 1287
Qy 227 GACCCCCACGAAGCCCTTGCGCGCGCGCCCTTATCGAATCTGCCGAACACAGCCTC 286
Db 1287 rThrProThr-----SerSerTh 1293
Qy 287 GATTTCGAATACATATTGGCGCAACGACATTTCCGGCAGGCTGCTGTTCACACCTCA-- 344
Db 1293 rValLeuSerThrThr-----ProLysLeuCysCysLeuTrpSerAs 1307
Qy 344 ----- 344
Db 1307 pTrIleAsnGluAspHisProSerSerGlySerAspGlyAspArgGluProPheAs 1327
Qy 345 -----TGTAACCTTCCGCGACAAACCGCGCGTTCGCGGTACGCC----- 380

```

```
Db      1327 pGlyValCysGlyAlaProGluAspIleGluCysArgSerValLysAspProHisLeuSe 1347
      QY      ||| |||:||||| |||
      380 -----
Db      1347 rLeuGluGlnHisGlyGlnLysValGlnCysAspValSerValGlyPheIleCysLysAs 1367
      QY      ||| |||:||||| |||
      380 -----
Db      1367 nGluAspGlnPheGlyAsnGlyProPheGlyLeuCysTyrAspTyrLysIleArgValAs 1387
      QY      -TGCTGTTGG-----
      381 -TGCTGTTGG-----ACGACAAAC 400
      1387 nCysCysTrpProMetAspLysCysIleThrThrProSerProProThrThrProSe 1407
      QY      ||| |||:||||| |||
      401 AGCGCGGGTGGAGCATCTCTCGCTCGACGACCATCCCAATATCGAAGTGGC 460
      1407 rProProThrThrThrThrLeuProProThrThrPro----- 1422
      QY      ||| |||:||||| |||
      461 CTGTTCAACCCCTCGTCTCTACGAAATGGCGCGCACTCGGCTACTGACCGACTTCCCC 520
      1423 ---SerProProThrThrThrThrThrThrProProProThrThrThrProSerProPr 1441
      QY      ||| |||:||||| |||
      521 CGCCTCAACCGCGCATGCACAAATCCTTTAGCGCGCAACCGCGCCACCATATCTC 580
      1441 oIleThrThr-----ThrThrProLeuProThrThrThrProSerPro----- 1456
      QY      ||| |||:||||| |||
      581 GCGCAGCGCAATATCGCGACGAATACTTCAAGTCGGTGAGGACACCGTTTTTCGCGCAG 640
      1457 -----ProIleSerThrThrThrThrProProProThrThrThrProSerProProTh 1474
      QY      ||| |||:||||| |||
      641 CTGGACATCTCGCACCGGAGCGTGTGGCGAAGTATCGCAGACTTCGAC----- 695
      1474 rThrThrProSerProProThrThrThrProSerProProThrThrThrThrThrPr 1494
      QY      ||| |||:||||| |||
      696 -----GCTACTGGCAAGCATTCGCGCCACACGCGCAGCGCATCATCGCGACGGC 748
      1494 oProProThrThrThrProSerProProMetThrThrProIleThrThrProAlaSerTh 1514
      QY      ||| |||:||||| |||
      749 AACATCGCGCAGGGTCTTCAAGCACTCGGATACACGACGAAACATCCAGACACGCGCTC 808
      1514 rThrThrLeu-----ProProThrThrThrThrProSerProProThrThrTh 1529
      QY      ||| |||:||||| |||
      809 CTGCGCTACCGGAAACCGTTCGAAC-----AGTCGCCCTCTACCAAAATACAGACG 862
      1529 rThrThrThrProProProThrThrThrProSerProProThrThrThrThrTh 1544
      QY      ||| |||:||||| |||
      863 GGACGATCGACTGCGAGAGCGTCCAAACCGCCTGATCAGCGACACCCCTGCAAAAGGA 922
      1545 -----ThrProIleThrProPr 1550
      QY      ||| |||:||||| |||
      923 CTCGACCGGACCGCGCAACCGCGGATTCGCGGAGGCTGCAAGACGCGCTCAAAACAG 982
      1550 oThrSerThrThr-----LeuPro-----ProThrThrThrProSe 1563
      QY      ||| |||:||||| |||
      983 CCGGAAAAAGGCTATCTGTTTCAACCTATTTCGTCCTTACAAAATCCGGACAGAC 1042
      1563 rProProProThrThrThrThrThrThrProProProThrThrThrProSerProThrTh 1583
      QY      ||| |||:||||| |||
      1043 GCATCGGCAAAACTGGTGAGGAGCGCATAGACGTTTACCGTCTCGACCA-----AC 1093
      1583 r-----ThrThrProSerProProThrThrThrThrThrThrThrTh 1593
      QY      ||| |||:||||| |||
      1094 TCGCTACAGCGGACCGAGTTCGCGCGTCCATTCGCGGTACGTCAAAAT----- 1142
      1593 rThrThrProProProThrThrThrProSerProProThrThrThrThrThrProPr 1613
      QY      ||| |||:||||| |||
      1143 -----ACGAAACCGCTGCTCAAAGCGGATCAAATC 1177
      1613 oProThrThrThrProSerProProThrThrThrThrProIleThrThrThrThrTh 1633
      QY      ||| |||:||||| |||
      1178 TACGAGCTGCAACCAACCATCGCTCCCGCGCACAAAGACAAAGGCGCTGACCGCGCAGC 1237
      QY      ||| |||:||||| |||

Db      1633 rThrLeuProProThrThrThrThrProSerProProProThrThr-----ThrThrTh 1650
      QY      TCCGTAAACCGCTGCATGCCAAAACCTTATTGTGGACGCGCAACGCGATCTTCATCGGC 1297
      1650 rProProProThrThrThrProSerProProThr-----ThrThrThrProSerProPr 1669
      QY      ||| |||:||||| |||
      1298 TCATTCAACCTCGACCGCCCGTTCGCGACGGCTCAATACCGAANTGGGGTGTCTATCGAA 1357
      1669 eThrThrThrThrProProProThrThrThrProSerProIleThrThrThrPr 1689
      QY      ||| |||:||||| |||
      1358 AGCCCAAAATCGAGAACAGATGGAGCGCACCTTCGCGCATACACACCCGGAATACGCC 1417
      1689 oSerProProThr-----ThrThrMetThrThrProSerProThrThrThrProSerSerPr 1708
      QY      ||| |||:||||| |||
      1418 TACCGCTTACCTCGACAAAACACACCGCTCGCAATGGCAGATCCCGCCACCCGCAAAA 1477
      1708 oIle-----ThrThrThrThrThrProSerSerThrThrThrProSerProPr 1725
      QY      ||| |||:||||| |||
      1478 ACCTACCGAAGAACCGCAAGCCAACTTTGGAAACGCATCGCGCGCAAAATCTATACC 1537
      1725 oThrThrMetThrThrProSerProThrThrThrProSerProThrThrThrMetTh 1745
      QY      ||| |||:||||| |||
      1538 CTGCTGCCCA 1547
      1745 rThrLeuPro 1748

RESULT 14
US-10-734-564-121
; Sequence 121, Application US/10734564
; Publication No. US20040157278A1
; GENERAL INFORMATION:
; APPLICANT: Christopher C Burgess et al
; TITLE OF INVENTION: Detection Methods Using TIMP1
; FILE REFERENCE: 1657/2012
; CURRENT APPLICATION NUMBER: US/10/734,564
; CURRENT FILING DATE: 2003-12-12
; NUMBER OF SEQ ID NOS: 138
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 121
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-734-564-121

Alignment Scores:
Pred. No.: 6,05e-14 Length: 5179
Score: 336.50 Matches: 141
Percent Similarity: 33.6% Conservative: 55
Best Local Similarity: 24.1% Mismatches: 198
Query Match: 11.8% Indels: 190
DB: 4 Gaps: 22

US-10-665-990A-13 (1-1561) x US-10-734-564-121 (1-5179)
QY 119 ACGGAAACCGCTGATTTCAAT-----ACTTCCAAACCTGTCTCTCTGGACACAC 166
Db 1248 ThrValGluLysHisPheAsnIleCysSerIleThrThrArgPro-SerThrLeuThrTh 1267
QY 167 ATCTGCAAAATCCGCGCACACCCCTCATAAACAGGGCTATCGACATCTACCTGCTCGAC 226
Db 1267 rPheThrThrIleThrLeuProThrThrProThrSerPheThrThrThrThrTh 1287
QY 227 GACCCCAAGAGCCCTTGGCGCGCGCCCTTATCGAATCTGCGGAACACAGACCTC 286
Db 1287 rThrProThr-----SerSerTh 1293
QY 287 GATTTCGAATACTACATTTGGCGCAACGACATTTCCGCGAGGCTGCTGTTCACACTCA-- 344
Db 1293 rValLeuSerThrThr-----ProlLysLeuCysCysLeuTrpSerAs 1307
QY 344 ----- 344
Db 1307 pTrpIleAsnGluAspHisProSerGlySerAspGlyAspArgGluProPheAs 1327
```

QY 345 -----TGATCTTCCGCGAGAACGCGCGTGCCTACGCC----- 380
Db 1327 pGlyValCysGlyAlaProGluAspIleGluCysArgSerValLysAspProHisLeuSe 1347
QY 380 ----- 380
Db 1347 rLeuGluGlnHisGlyGlnLysValGlnCysAspValSerValGlyPheIleCysLysAs 1367
QY 380 ----- 380
Db 1367 nGluAspGlnPheGlyAsnGlyProPheGlyLeuCysTyrAspTyrIleValAs 1387
QY 381 -TGCTGTGG-----ACGAAACAAC 400
Db 1387 nCysCysTrpProMetAspLysCysIleThrThrProSerProProThrThrProSe 1407
QY 401 ACGGCGGGTGGAGGATCTCTGCTCGCCCTCGACAGCCATCCCATATCGAAGTGC 460
Db 1407 rProProProThrThrThrLeuProProThrThrPro----- 1422
QY 461 CTGTTCAACCCCTCGTCTAGCAAAATGGCGGCGACTCGGCTACCTGACCGACTTCCC 520
Db 1423 -----SerProProThrThrThrThrThrProProProThrThrThrProSerProPr 1441
QY 521 CGCCTCAACCGCGCATGCACAACAATCTTTACCGCGGACACCGCGCCACCACATCTC 580
Db 1441 oIleThrThr-----ThrThrThrProLeuProThrThrThrProSerPro----- 1456
QY 581 GCGGAGCAATATCGGCGGAGCAATACTTCAAGTCGGTAGGACACGCTTTTCGCGAC 640
Db 1457 -----ProIleSerThrThrThrThrProProProThrThrThrProSerProProTh 1474
QY 641 CTGACATCTCGCCACCGGCGAGTGTCTGCGGAGTATCGCAGGACTTCGACC----- 695
Db 1474 rThrThrProSerProProThrThrThrProSerProProThrThrThrThrThrPr 1494
QY 696 -----GCTACTGGGAAGCATTCGCGCCACACCGCCGCGCATCATCCGCGCGG 748
Db 1494 oProProThrThrThrProSerProProMetThrThrThrProIleThrProProIleSerTh 1514
QY 749 AAGATCGGAAGGTCTTCAAGCTCGGATACAAACGACGAAACATCGACAGCGGCTC 808
Db 1514 rThrThrLeu-----ProProThrThrThrProSerProProThrThrTh 1529
QY 809 CTGCGCTACCGGAAACGTCGAC-----AGTCGCGCTCTACCAAAATAACACAG 862
Db 1529 rThrThrThrProProProThrThrThrProSerProProThrThr----- 1544
QY 863 GGAGCATCGACTGGCAGCGTCCAAACCGCGCTGATCAGCGACACCCCTGCAAAAGGA 922
Db 1545 -----ThrProIleThrProPr 1550
QY 923 CTCGACCGGACCGCGGAAACCGCGATTCGCGGAGGTGCGAAGACGCGCTCAACAG 982
Db 1550 oThrSerThrThr-----LeuPro-----ProThrThrThrProSe 1563
QY 983 CCGGAAAAAGCGTATCTGGTTTACCTATTTCCTCCCTACAAATCGGCGACAG 1042
Db 1563 rProProProThrThrThrThrThrProProProThrThrThrProSerProProThrTh 1583
QY 1043 GCACTGGCAAAACTGGTGCAGGCGCATAGACGTTTACCGTCTCTGACCA-----AC 1093
Db 1583 r-----ThrThrProSerProProThrThrThrThr 1593
QY 1094 TCGCTACAGGGACCGAGTTCGCGCGTCCATTCGCGTACGTCAAAAT----- 1142
Db 1593 rThrThrProProProThrThrThrProSerProProThrThrThrThrThrProPr 1613
QY 1143 -----ACGAAACCGTGTCTCAAGCGCGCATCAACTC 1177
Db 1613 oProThrThrThrProSerProProThrThrThrProIleThrProProThrThrThrTh 1633

QY 1178 TACGAGTGCAAACCAACCATGCGTCCCGCCACAAAAGAGCGCTGACCGGCAGC 1237
Db 1633 rThrLeuProProThrThrThrProSerProProProThrThr-----ThrThrTh 1650
QY 1238 TCCGTAAACGAGCTGCATGCAAAACCTTTCATTTGGAGCGCAACGCACTTTCATCGGC 1297
Db 1650 rProProProThrThrThrProSerProProThr-----ThrThrThrProSerProPr 1669
QY 1298 TCATTCAACCTCGACCCCGTTCCGACAGGCTCAATACCGAAATGGGGTCTGTCATCGAA 1357
Db 1669 eThrThrThrThrThrProProProThrThrThrProSerProIleThrThrThrPr 1689
QY 1358 AGCCCCAAAATCGAGAACAGATGAGCGACCGCTCCCGCATACACACCGCAATACGCC 1417
Db 1689 oSerProProThr-----ThrThrMetThrThrProSerProThrThrThrProSerPr 1708
QY 1418 TACCGCTTACCTCGACAAACACACCGCTGCAATGGCACGATCCCGCACCGCAAAA 1477
Db 1708 oIle-----ThrThrThrThrProSerSerThrThrThrProSerProPr 1725
QY 1478 ACCTACCGAACGACCGAAGCAAACTTTGGAAACGATCGCGCGCAAAATCCTATCC 1537
Db 1725 oThrThrMetThrThrProSerProThrThrThrProSerProThrThrThrMetTh 1745
QY 1538 CTGCTGCCCA 1547
Db 1745 rThrLeuPro 1748
RESULT 15
US-09-801-368-108
; Sequence 108, Application US/09801368
; Patent No. US20020128250A1
; GENERAL INFORMATION:
; APPLICANT: Busby, Robert
; APPLICANT: Cali, Brian
; APPLICANT: Hecht, Peter
; APPLICANT: Holtzman, Doug
; APPLICANT: Madden, Kevin
; APPLICANT: Maxon, Mary
; APPLICANT: Milne, Todd
; APPLICANT: No. US20020128250A1man, Thea
; APPLICANT: Royer, John
; APPLICANT: Salama, Sofie
; APPLICANT: Sherman, Amir
; APPLICANT: Silva, Jeff
; APPLICANT: Summers, Eric
; TITLE OF INVENTION: Methods for Improving Secondary Metabolite Production in Fungi
; FILE REFERENCE: 109272.147
; CURRENT APPLICATION NUMBER: US/09/801,368
; CURRENT FILING DATE: 2001-03-07
; PRIOR APPLICATION NUMBER: US 09/487,558
; PRIOR FILING DATE: 2000-01-19
; PRIOR APPLICATION NUMBER: US 60/160,587
; PRIOR FILING DATE: 1999-10-20
; NUMBER OF SEQ ID NOS: 440
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 108
; LENGTH: 1367
; TYPE: PR1
; ORGANISM: Saccharomyces cerevisiae
US-09-801-368-108
Alignment Scores:
Pred. No.: 3,62e-13 Length: 1367
Score: 323.50 Matches: 147
Percent Similarity: 40.7% Conservative: 88
Best Local Similarity: 25.4% Mismatches: 234
Query Match: 11.3% Indels: 109
DB: 3 Gaps: 18
US-10-665-990A-13 (1-1561) x US-09-801-368-108 (1-1367)
QY 80 TGTTCATGTTCTTCATGTTGCCCCCACTG-----GAAGAAGCGAGC 121


```
QY 1146 -----GAAACCGTGCTCAAGCCGGCATCA 1172
Db 348 ProProIleSerSerThrIleSerValGlnProSerSerSerSerProThrThr 367
QY 1173 AACTCTACGAGTGCACCAACCATCGCTCCCGCCACAAAAGACAAAGCCGTGACCG 1232
Db 369 SerThrThrSerValGlnProSerSerSerGlySerAlaProThrThrSerAlaThrSer 387
QY 1233 GCAGCTCCGTAAACCGCTGCATGCAAAACCTTCATTGTGACGGCAAAACGATCTTCA 1292
Db 388 ValGlnPro-----SerSerSerSerSerValProThrThrSerAlaThrSer 403
QY 1293 -----TCGGCTCATTCAACCTGCAGCCCGCTCGCAGCGGTCAATACCGAAATGGGG 1346
Db 404 ValArgSerSerSerSerThrProIleProThrThrThrSerValGlnProSer 423
QY 1347 TCGTCATCGAAGCCCAAAATCG-----CAGAACAGATGGAGCGCACCC 1391
Db 424 SerSerSerSerValProThrThrSerAlaThrSerValGlnThrSerSerSerSer 443
QY 1392 TCGCGATACCA-----CACCGGAATACGCTTACCGGTTCACCTCGACA 1436
Db 444 ThrProIleProSerThrThrSerValGlnProSerSerSerSerSerAlaProThrThr 463
QY 1437 AACACAAACCGCTGC-----AATGGCAGATCCCGCACCCGAAACCTACCGAAG 1490
Db 464 SerAlaThrSerValGlnProSerSerSerSerProIleSerSerThr---Ile 482
QY 1491 AACCCGAAGCCAAACTTTGGAACGATCGCCGCAAAATCC 1532
Db 483 SerValGlnProSerSerSerSerSerProThrThrSer 496

RESULT 21
US-10-437-963-125004
; Sequence 125004, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 125004
; LENGTH: 497
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(497)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_27690C.1.pap
US-10-437-963-125004

Alignment Scores:
Pred. No.: 1,57e-11 Length: 497
Score: 298.00 Matches: 145
Percent Similarity: 37.7% Conservative: 56
Best Local Similarity: 27.2% Mismatches: 170
Query Match: 10.4% Indels: 162
DB: 4 Gaps: 29

US-10-665-990A-13 (1-1561) x US-10-437-963-125004 (1-497)
```

```
QY 267 AATCTCGCAACACAGCCTCGATTGC-----AATACTACA 302
Db 1 AsnLeuHisAsnAsn---AsnLeuCysGlySerValProValArgGluGlyAsnLysThr 19
QY 303 TTGGCGCAACGACATTTCCGCGCAGG-----TGCTGT-----TCAACC 341
Db 20 GlyProSerSerSerGluValAlaGlySerSerCysCysAspGluAspGluAlaAspLeu 39
QY 342 TCATGTACTTTGCCGACAGACGCGGTACCGCTGCTGCTGTGTGGACGACAAACA 401
Db 40 SerCysLeuMetProThrAlaAlaArgAlaSerAlaAlaAala---AlaValThrAla 58
QY 402 CGCGCGGTGTGACGATCTCTGCTCGCCCTCGACAGCATCCCAATATCGAAG-----455
Db 59 ArgSerSerSerSerSerAlaGlyProAlaThrSerAlaThrSerSerGlyArg 78
QY 456 -----TGCGCTGTTCAACCCCTTCGTCCTAC 482
Db 79 SerArgGlyAlaProProAlaSerTrpArgProAlaProCysSerThrProProSer*** 98
QY 483 GCAATGGCGGCACACTCGGCTACCTGACCGACTTCCCCCGCTCAACCGCGCATGCACA 542
Db 99 SerAlaAlaProTrpArgAla-----ProThrThrProThrSerAlaArgSerSerThr 116
QY 543 ACA-----AATCTTTACCGCGCACACCGCGCCA 572
Db 117 SerArgArgProAlaProSerArgProArgGlySerProProProProThrSerArgArg 136
QY 573 CCATACTCGCGGACGCA-----ATATCGCGCAGC 602
Db 137 ProThrSerSerSerSerSerArgAlaSerProSerProProProSerAlaSer 156
QY 603 AATACTCAAAGTCGGTAGGACACCGTTTTCGCGCAGCTGGACATCTCGCCACCGGCA 662
Db 157 AlaProSerProAlaThrGlyThrProProSerProProThrThrSerSerSerProThr 176
QY 663 GCG-----TCGTGCGCGAAGTATCGCAGACT 689
Db 177 AlaProGlyProProProArgAlaThrArgSerCysSerSerSerSerArgArgThrThr 196
QY 690 TCGACCGCTACTGGGCAAGCCATTCGCGCCACACGCGCACGCGCATCTCCGCGCGCA 749
Db 197 AlaThrPro-----ProProSerThrProAlaArgSerAlaAlaAlaSer 211
QY 750 ACATCGCAAGGTTCTCAAGCATTCCGGATACACAGCAAAACATCCAGACACGCGCTCC 809
Db 212 ArgArgProProHisAlaSerSer***ProSerThrThrSerAlaProSerThrProSer 231
QY 810 TGC-----GCTACCGCGAAACCGTCGAACAGTCGCGCCCTCTACCAAAAAA 854
Db 232 CysArgProArgArgAlaGlyAlaSerProProAlaGlyAlaProAlaProGlyGlu 251
QY 855 TACAGACGGGACGATCGACTGGCAGAGCGTCCAAACCCGCTGATCGACGACACCCCTG 914
Db 252 CysArgArg-----ProArgProAla---SerSerGlnProCys 263
QY 915 CAAAG---GACTCGACCGGACCGCCGCAACCCCGGATTCGCGGAGGCTGCAAGACG 971
Db 264 ArgArgCysSerSerThrAlaThrSerThr-----GlyCys-----275
QY 972 CGCTCAACAGCCCGAAAAAGCGTCTATCTGTTTTCACCCCTATTTCGTCCTTACAAAT 1031
Db 276 ---SerSerGlnProProProAlaThr-----SerArgSerGlyGluSerThrArgSer 292
QY 1032 CCGGCACAGACGCACTGGCAAACTGGTGCAGGACGCATAGACGTTTACCG-----TCC 1085
Db 293 ProSerAlaPro-----SerSerSerGlyArgSerLeuProThrThrSer 307
QY 1086 TGACCAACTCGTACAGGGA-----CCGACGTTGCGCCGCTCCATTCGCGTAGCTCA 1139
Db 308 SerSerThrArgSerSerArgAlaSerProProProProSerSerAlaProSer 327
```



```
QY 1140 AATACGAAACCGCTGCTCAAAGCGGCATCAAACTCTACGAGCTGCAACCAACCAACCATG 1199
Db   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
328 ProAlaThrGly-----ThrProProSerProProThrThrSerSerSerProThrAla 345
QY 1200 CGTCCCGCCA-----CAAAGCAAAAG 1223
Db   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
346 ProGlyProProProAlaAlaThrArgSerPheSerSerArgArgArgAsnPro 365
QY 1224 GCTGACCGGAGCTCCGTAACGAGCTGCATGCGCAAAACCTTCATTGTCGACGCAAAAC 1283
Db   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
366 AlaGluProAlaAlaProPro-----SerThr 375
QY 1284 GAATCTCATCGCTCATTTCAACCTCGACCCCGCTTCGACGGCTCAATACCGAAATGG 1343
Db   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
376 ProAlaArgSerAlaAlaAlaArgArgProProProAlaSerSerSerProSerThr 395
QY 1344 CGTCTCATCGAAAGCCCCAAATCGCAGAACATGAGGACCGCTCG---CCGATA 1400
Db   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
396 Ala-----SerAlaProSerThrProSer 403
QY 1401 CCACACCGAATACGCTACGCGTTACCTCGACAAACACACCGCTGCAATGGC--- 1457
Db   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
404 ArgArgProGlyArgAlaGlyAlaSerProSerSerProThrThrAlaAlaSerArgGlyIle 423
QY 1458 -----ACGATCCCGCCA-----CCGAAACAACTACC----- 1484
Db   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
424 ThrSerSerThrCysProProAlaIleThrSerArgProThrSerSerProAlaCysHis 443
QY 1485 -----CGAACGAACCGAAGCCAACTTTGGA 1511
Db   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
444 ArgCysSerSerThrAlaThrSerThrGlyCysSerSerGlnProProProAlaThrGly 463
QY 1512 AACGATCGCGCGCAAAATCCTATCCCTGCTGCCATCG 1550
Db   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
464 ArgSerGlyGlySerCysArgSerProSerAlaProSer 476
```

RESULT 22

```
US-10-084-846A-5
; Sequence 5, Application US/10084846A
; Publication No. US2004006026A1
; GENERAL INFORMATION:
; APPLICANT: WEITNAUER, GABRIELE
; APPLICANT: TREFFZER, AXEL
; APPLICANT: BECHTHOLD, ANDREAS
; TITLE OF INVENTION: AVILAMYCIN DERIVATIVES
; FILE REFERENCE: 1974-005
; CURRENT APPLICATION NUMBER: US/10/084,846A
; CURRENT FILING DATE: 2003-02-25
; PRIOR FILING DATE: PCT/EP01/09815
; PRIOR FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: DE 101 09 166.4
; PRIOR FILING DATE: 2001-02-25
; NUMBER OF SEQ ID NOS: 120
; SOFTWARE: PatentIn Ver. 3.2
; SEQ ID NO 5
; LENGTH: 19723
; TYPE: PRT
; ORGANISM: Streptomyces viridochromogenes
; FEATURE:
; OTHER INFORMATION: Protein 3: amino acid sequence encoded by coding strand 1.
; OTHER INFORMATION: Start codon: atc, Start position: nucleotide 3.
US-10-084-846A-5
```

Alignment Scores:

Pred. No.:	3.11e-11	Length:	19723
Score:	297.50	Matches:	158
Percent Similarity:	37.0%	Conservative:	63
Best Local Similarity:	26.5%	Mismatches:	212
Query Match:	10.4%	Indels:	165
DB:	4	Gaps:	29

US-10-665-990A-13 (1-1561) x US-10-084-846A-5 (1-19723)

```
QY 100 GCCCCCACTGGAAGAACGGACGGAAAGCCGTCATTTCAATTAATCTCCAAACACCTGCTCCTCT 159
Db   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
18236 AlaGlyThrGlyArgser-----ProAla 18243
QY 160 GGACAAACATCTGCTGAAAATCCGGCACACCCCTCATAAACAACGCGGTATCCGACATCTACT 219
Db   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
18244 CysSerArgProArgGlnProThrCysProSer---CysArgGlyValArgProAlaPro 18262
QY 220 GCTCGA-----CGACCCCAACGAAAGCCTTGCCGC-----CCGCGC 255
Db   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
18263 GlyArgValArgAlaAlaGlyAlaAlaSerProValArgSerThrArgTrpSerProArg 18282
QY 256 CGCCCTTATCGAATCTCCG---AACACAGCCCTCG-----ATTGCAATACTACATTT 305
Db   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
18283 ArgArgGlyProLeuCysProGlyAsnSerAlaSerProProCysGlyArg----- 18300
QY 306 GCGCAACACGACATTTCCGGCAGGCTGCTGTTCAACCTCATGT-----ACCTTGCCGCAG 359
Db   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
18301 -----ProAlaSerAlaCysProProSerAlaTrpSerThrGlySerSer 18315
QY 360 AACCGCGGCTGCGGTACGCTGCTGTGGACGACAACAACGCGCGGTGGACGATC 419
Db   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
18316 SerAlaAlaAlaSerTrpSerCysTrpProAsnSerProHisAlaProIleArgThr 18335
QY 420 TCCTGC-----TCGCCCTCGACAGCCATCCCAATA-----TCGAAG 455
Db   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
18336 ArgCysArgProValGlnAsnProSerThrGlyAlaArgThrAspLysSerGlnSerArg 18355
QY 456 TGCCTCTGTTCAACCCCTTCGCTCAGCAATGGCGCCTACGCTACCTGACCCGACT 515
Db   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
18356 CysAlaThrPheValProThrThrGlnThrAsnSer-----ProThr 18369
QY 516 TCCCCCGCTCAACCGCGGATGATGCAACAATCTTTACCG---CCGACAAACCGCGCA 572
Db   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
18370 SerAlaThrSerSerThrHisAlaThrAlaAspProSerProLeuProAlaThrPro 18389
QY 573 CCATACCTCGCGGACGCAATATGGCGACGAATACT-----TCAAGTCG 617
Db   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
18390 GlyHisSerThrGlyAlaSerProThrGlyThrGlyProThrThrTyrSerSerAla 18409
QY 618 GTGAGGACACCGTTTTCGCGACCTGGACATCTCCGACCGCA----- 662
Db   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
18410 AlaSerTrpArgProThrProSerSerThrAlaSerArgProAlaMetValSerPhePhe 18429
QY 663 -----GCGTCGTCGGCGAAGTATCGACAGCTTCGACGCTTCTGCGGCA 707
Db   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
18430 AlaSerAlaMetThrAlaThrSerSerAlaSerArgSerThrThrAlaGluArgAlaTyr 18449
QY 708 GCCATTCCGCGCCACACGCCA---CGCGCATCATCCGACGGGCAACATCGGCAAGGTC 764
Db   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
18450 ProGluProProThrThrProThrLysAlaAlaValCysSerTrpSerAlaArg--- 18468
QY 765 TTCAAGCACTCGGATACACGACGAAACATCCAGACACGCGCTCCTGCGCTACCGCAAA 824
Db   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
18469 -----TrpProThrSerGlyAlaTrpAlaSerGluArg 18479
QY 825 CGGTGCAACAGTCGCCCTCTACCAAAAATACAGA-----CGG 863
Db   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
18480 ProAlaLysSerSerGlyProSerSerProCysArgAlaLeuValGlnPheLeuTrpArg 18499
QY 864 GACGCATCGACTGCGACAGCGT-----CCAAACCCGCC 896
Db   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
18500 ProGlnSerPheGlyArg-ThrGlyGluAlaGlnValAspMetProLeuProThrProPr 18519
QY 897 TGATCAGCGACACCCCTGCAAAAGGACTCGA----- 927
Db   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
18519 o-----ArgArgProHisLysArgThrThrAspGlyGlyProLeuArgTrpGlnCysVa 18537
QY 927 ----- 927
Db   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
18537 lIleProGlyGluAspTrpProProSerAlaArgProProArgProSerProArgAlase 18557
```

```
QY 928 -----CCGCGACCCGCGCAAAACCCCGGATTCGCCGGGAGGCTGCAAGACG 971
Db 18557 tSerProLeuSerProProArgProProArgAlaValSerSerThrArgAlaArgG1 18577
QY 972 C-----GCTCAAAACAGCCGAAAAAGCGTCTATCTGCTTCCACCTATTTCG 1019
Db 18577 nThrProTrpProThrAlaThrSerSerArgThrThrProValArgValProAr 18597
QY 1020 TCCCTACAAAATCCGGCAGACGACCTGGCAAAACTGGTGAGGACGCGCATAGACGTTA 1079
Db 18597 gProAspAlaProSerSerArgSerThrAlaProThrSerSerAlaAaGThrArgArg-- 18616
QY 1080 CGGTCTGACCAACTCGCTACAGGCGACGCGGCTGCGCGCTCATCTCCGGCTACGTCA 1139
Db 18617 -----ProThrAlaThrArgSerArgProProArgArg-----ArgProArgAr 18631
QY 1140 AATACCGAAAAACCGTGTCAAAAGCGGCATCAAACTCTACGAGCTGCAAC----- 1190
Db 18631 gSerProProThrArg--ArgSerProProTrpSerArgThrAlaCysSerAlaSerThr 18650
QY 1191 -----CCAACCATGCCG-----TCCCGG 1208
Db 18651 ArgProArgProThrArgProProGlyAlaTtpThrAlaTtpThrArgArgThrSerPro 18670
QY 1209 CCACAAAAGACAAAGCCCTGACCGCAGCTCCGTAAACGAGCTGATGCCAAAACCTTCA 1268
Db 18671 ThrAlaAaGThr-----ProThrArgThrArgProGlyArgGluArgArgThrSer 18687
QY 1269 TTGTGGAGCGCAACGATCTTCATCGGCTCATTCACCTGACCCCGTTCGCGCAGCGC 1328
Db 18698 SerThrProAlaSerAlaSerProThrAlaThrSerAlaAaGlyPro---ProThrAla 18706
QY 1329 TCAATACCGAAATGGCGTGTCTATCGAAGGCCCAAAATCCGACAAACAGATGGAGCGCA 1388
Db 18707 ThrThrPro-----SerThrThrThrProProArgThrAlaThrAlaThrAlaArg 18724
QY 1389 CCTCGCGGATACCAACCGCAATAGCTACCGGTGTACCTCGACAAACACACACCGCC 1448
Db 18725 ThrTtpProAlaAaGSerProAlaThrProThrAlaSerProArgProArgSerAla 18744
QY 1449 TGCAATGGACCATCCGCGCCACCCGAAACAACTTACCAGAAACCCGAAAGCCAAACTT 1508
Db 18745 SerAlaCysThrThrProAlaArgAlaProProProArgSerSerProAlaSerThrGly 18764
QY 1509 GGAACGATCGCGC---CAAAAATCTTATCTCCTGCTGCCCATCG 1550
Db 18765 SerProGlyThrProSerSerArgProSerProThrCysProSer 18779

RESULT 23
US-10-418-861B-59
; Sequence 59, Application US/1041861B
; Publication No. US20040010131A1
; GENERAL INFORMATION:
; APPLICANT: da Silva, Ana Claudia Rasera
; APPLICANT: Farah, Shaker Chuck
; APPLICANT: Queglio, Ronaldo Bento
; APPLICANT: Reinach, Fernando de Castro
; APPLICANT: Ferro, Jesus Aparecido
; APPLICANT: De Oliveira, Julio Cezar Franco
; APPLICANT: De Laia, Marcelo Luiz
; APPLICANT: Setubal Joao C.
; APPLICANT: Furlan, Luiz Roberto
; TITLE OF INVENTION: Isolated Xanthomonas nucleic acid molecules, proteins encoded the
; FILE OF INVENTION: uses thereof
; FILE REFERENCE: FAPESP 205.1 US
; CURRENT APPLICATION NUMBER: US/10/418,861B
; CURRENT FILING DATE: 2003-04-17
; PRIOR APPLICATION NUMBER: US 60/374,620
; PRIOR FILING DATE: 2002-04-22
; NUMBER OF SEQ ID NOS: 85
; SEQ ID NO 59
; LENGTH: 486
; TYPE: PRT
```

```
; ORGANISM: Xanthomonas
; FEATURE:
US-10-418-861B-59

Alignment Scores:
Pred. No.: 2,13e-11 Length: 486
Score: 296.00 Matches: 123
Percent Similarity: 38.0% Conservative: 63
Best Local Similarity: 25.2% Mismatches: 209
Query Match: 10.4% Indels: 94
DB: 4 Gaps: 14

US-10-665-990A-13 (1-1561) x US-10-418-861B-59 (1-486)
QY 83 TCATGTTCTTCATGGTTGCCCTCCCACTGGGAAGAACGGACGGAAGCCGTCATTTCAATACT 142
Db 89 SerTyrSerValCysProAspAlaAspCysThrGlu----- 102
QY 143 TCCAAACCTGTCTCTCGACAAACATCTCTGCAAAATCCGGCACACCCCTCATAAACAACGGG 202
Db 103 -----LeuAlaLysIleAlaGlnSerThrThrGlyLeuAlaProSerSer 117
QY 203 CTATCCGACATCTACCTCTCGACGACCCCCACGAAGCCCTTGGCGCGCGCGCCCTT 262
Db 118 AlaThrGluValHisTtpLeuValAspGlyAlaAlaThrTyrAlaAlaIleleGluAla 137
QY 263 ATCGAATCTGCGCAACACAGCCTCGATTGCAATCTACTATTGGCGCAACGACATTTCC 322
Db 138 IleArgGlyAlaAaGAspHisIleHisLeuGluTyrTyrIlePheGlnProAspHisSer 157
QY 323 GGCAGGCTGTGTTCACCTCATCTTCCCGCAGAACCGCGCGTGGCGGTACGCGCTG 382
Db 158 GlyThrAlaIleCysAlaAlaLeuMetGluArgAlaGlyValLysValArgLeu 177
QY 383 CTGTTGGACGACAAACACACGCGGGTTGACGACATCTCTCTCGCCCTCGACACCCAT 442
Db 178 LeuMetAspAlaIleGlySerSerAlaMetThrArgArgAlaLeuArgThrLeuArgGlu 197
QY 443 CCCAATATCGAAGTGGCGCTTTCAAACCCCTTGTCTCTACGCAAAATGGCGGACATCGGC 502
Db 198 AlaGlyValGluThrAlaTtpPheHisProSerGlnLeu----- 210
QY 503 TACTGTACCGGACTTC-----CCCGCCTCAACCGCGCATGCACAAACATCTTTACC 556
Db 211 ---LeuLysProPheLysArgProTtpLeuAsnLeuArgThrHisArgLysValIleVal 229
QY 557 GCCGACAAACCGCGCCACCATATCTCGCGGACGCAATATCGCGACGAATACTTCAAAGTC 616
Db 230 IleAspGlyArgIleGlyPheThrGlyGlyIleAsnValThrAspAspGluAsnGln 249
QY 617 GGTGAGGACACGGTTTTCGCGACCTTGACATCTCTCGCCACCGGACGCGTCTCGCGGAA 676
Db 250 ValArgLysGlnAlaTyrArgAspLeuHisValArgLeuGlnGlyHisValValArgSer 269
QY 677 GTATCGCACGACTTCGACCGCTACTTGGGCAAGCCATTCGCCCAACCGCACCGCGCATC 736
Db 270 LeuGlnLeuValPheLeuGluAspTtpLeu-----TyrAlaThrSerGln 284
QY 737 ATCCGACGCGGCAACATCGGCAAGGCTTTCGAAGCACTCGGATACAAACGACGAAATATCC 796
Db 285 GlyArgAlaAlaPheHisGlyGlnLeu-----TtpProAspAspMetPro 300
QY 797 AGACACGGCTCTCGGCTACCGGAAACCGTTCGAAACAGTCCGCCCTCTACCAAAAAATA 856
Db 301 ThrArgAla-----GlnGlyThrValAspAlaGlnValLeu-----Val 313
QY 857 CAGACGGACGATCGACTGCGCAGAGCGTCCAAACCCCGCTGATCAGGACGACACCCCTGCA 916
Db 314 SerGlyProAspSerSerTtpGluAlaIleHisArgLeuMetValAla----- 329
QY 917 AAAGGACTCGACCGCGACCGCGCAACCCGCGGATTCGCGGAGGCTGCAAGACGCGTTC 976
Db 330 -----AlaIle 331
```



```
Db      322 rThrAsnSerProArgSerThrSerValGlyArgSerProSerProCysProSerLe 342
QY      1127 TCCGGCTAGTCAATACCGAAACCGGTGCTCAAGCGCGCATCAAACTCTACAGGTG 1186
Db      342 u-----SerSerProCysSerProProSerProSer----- 353
QY      1187 CAACCAACCATCGCGTCCCGCCACAAAAGAGCGCTGACCGGAGCTCCGTAACC 1246
Db      354 -----ProSerProSerGlnSerSerThrProSerSerSerAlaProSerPr 369
QY      1247 AGCTGCGATGCCAAACCTTCATTGTGGACGGCAACAGCATCTTCATCGGCTCATCAAC 1306
Db      369 o-----**AlaSerAlaSerAlaSerThrLysSerAl 380
QY      1307 CTCGACCCCGTTCCGACGGCTCAATACCGAAATGGGGCGTCTATCGAAAGCCCAAA 1366
Db      380 aSerValIleArgGlyGly-----ValSerSerAlaProSe 393
QY      1367 ATCGCAGAACAGATGGAGCGCACCTCGCCGATACACACCGCAATACGCT----- 1418
Db      393 r-----SerAlaProSerProProProProSerSerSerProProSe 408
QY      1419 -----ACGCGTTACCTCTCGACAAACACACCGCGCTCAATGGCAC 1459
Db      408 rSerSerAlaArgArgSerProSerProThrArgArgTrpThrSerCysSerGlyTh 428
QY      1460 GATCCCGCCACCGGAAACCTACCGACGACCGGAGGCA---AATTTGGAAACGC 1516
Db      428 rAlaPro-----ProThrThrSerArgProThrProAlaSerSerSerAl 444
QY      1517 ATCGCGCAAAATCTATCTCTGCTGC 1544
Db      444 aSerProProProThrArgArgArgCys 453
```

RESULT 26

```
US-10-437-963-176617
; Sequence 176617, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 176617
; LENGTH: 533
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(533)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_74349C.1.pep
US-10-437-963-176617
```

Alignment Scores:

Pred. NO.:	3,72e-11	Length:	533
Score:	292.50	Matches:	139
Percent Similarity:	34.3%	Conservative:	39
Best Local Similarity:	26.8%	Mismatches:	179
Query Match:	10.3%	Indels:	163
DB:	4	Gaps:	24

```
US-10-665-990A-13 (1-1561) x US-10-437-963-176617 (1-533)
QY      154 CCTCTGGACAAACATCTCTGCAAAATCCGGCACACCCCTCATAAACAACGGGCTATCCGACAT 213
Db      7 ProProThrProIleProProProValThrThrThrThrSerCysArgArg 26
QY      214 CTACTGCTCGACGACCCCGACGAAGCGCTTGGCCCGCGCGCCCTTATCGATCTGCG 273
Db      27 SerAlaAla***ValProProProAspProThrGluGlyArgHisHisHis-----Cys 44
QY      274 CGAACACGCTCGATTGTGCAATACTACATTTGGCGCAACGACATTTCCGCGAGCGTCT 333
Db      45 ArgSerProProLeuAla-----AlaArg 53
QY      334 GTTCAACCTCATGTACTTGTCCGCGAAGCGGGGT-----GCGGTACGGCT 381
Db      54 ProProProHisAlaProLeuGlnIleArgArgArgGluGlyAlaAlaAlaSerArgPro 73
QY      382 GCTGTTGGACGACAAACAACGCGCGGTGTGGACGATCTCTGCTGCGCCCTCGACAGCA 441
Db      74 AlaAlaGly-----SerGlyGlyArgGluGlyValAlaProSerProProLeuPro 91
QY      442 TCCCAATATCGAAGTGGCGCTGTTCAAACCCCTTCTCTACGCAAAATGGCGCGCACT--- 498
Db      92 --ProLeu***AlaThrAlaAlaSerThrProProAspLeuAlaGlu-AlaArgThrGly 110
QY      499 CGGCTACTGACGACGACTTCCCGCGCTCAACCGCGCATGCAACAACAAAT----- 548
Db      111 ProProProProArgProProProProProHisLeuProSerProThrProIleLeu 130
QY      549 CTTTACCGCGCACAAAC-----GCGCCACCATACTCGCGGAGCGCATATCG 596
Db      131 ProProProProThrThrHisProGlyGlyProProProThrHisLeuProAlaLeuSer 150
QY      597 GCGACGAATATCTCAAAAGTCGGTGAGGACACCGTTTTTCGCG-----ACC 641
Db      151 ProLeuLeuSerSerPheSer-----ProProLeuSerProLeuProGlnLeu**Ser 168
QY      642 TGGACATCTCG---CCACGGGAGCGTCTGTCGGCGAAGTATCCGACGACTTCGACCGCT 698
Db      169 LeuThrSerSerArgProProAlaSerThrSerLeuLeuLeuLeuLeuLeuLeuSer----- 186
QY      699 ACTGGCAAGCATCTCGCGCCACAAACGCGCGCATCATCCGCGCGGCAACATCGGCA 758
Db      187 -----HisValProProPro***Pro----- 193
QY      759 AGGTCTTCAAGCACTCGGATACAAACGACGAAACATCCAGACACGCGCTCTCTGCGCTACC 818
Db      194 -----ProProProProAspArgArg-----LeuPro 202
QY      819 GCGNAACCGTCGAACAGTCGCGCTCTACCAAAAATACAGACGGGACGATCGACTGGC 878
Db      203 ProProProThrProArgArgProLeuGlyHis----- 213
QY      879 AGAGCGTCCAAACCGCGCTCATGAGGACACCCCTGCAAAAGGACTCGACCGCGACCGCC 938
Db      214 ArgThrProAspProAlaAlaAsnAlaSerProIleAsp----- 226
QY      939 GCAACCGCGGATTTCGCGGAGGCTGCAAGACGCGCTCAAAACAGCCGCAAAAAGCGCTCT 998
Db      227 -----CysGlnThr----- 229
QY      999 ATCTGTTTCAACCTATTTCTGCTCAAAAATCCGGCACAGACGCACTGGCAAAACTGG 1058
Db      230 -----ProProLeuProProProProAlaSerThrArgSerGlyAsn--- 245
QY      1059 TGCAGGACGCATAGACGTTACCGTCTCGACCAACTCGCTACAGGCGCGACGCTTCGCG 1118
Db      246 -----ValAspProCysPro 250
QY      1119 CCGTCCATTCCGGCTACGTCAAAATACGAAAAACCGGTCTCAAAGCGCGCATCAACTCT 1178
Db      ::::||||| :||| :::|
```


QY 1532 CTATCC 1537
Db 489 ValSer 490

RESULT 29
US-09-966-521-18
; Sequence 18, Application US/09966521
; Publication No. US20030087321A1
; GENERAL INFORMATION:
; APPLICANT: TOMICH, Che-Shen
; APPLICANT: QUINN, Cheryl
; APPLICANT: ARVIDSON, Staifan
; APPLICANT: HARRIS, Douglas
; APPLICANT: MOTT, John
; TITLE OF INVENTION: ANTIMICROBIAL METHODS AND MATERIALS
; CURRENT APPLICATION NUMBER: US/09/966,521
; CURRENT FILING DATE: 2001-09-28
; NUMBER OF SEQ ID NOS: 145
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 18
; LENGTH: 493
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-966-521-18

Alignment Scores:
Pred. No.: 9,27e-11 Length: 493
Score: 286.50 Matches: 106
Percent Similarity: 38.2% Conservative: 86
Best Local Similarity: 21.1% Mismatches: 169
Query Match: 10.0% Indels: 141
DB: 3 Gaps: 16

US-10-665-990A-13 (1-1561) x US-09-966-521-18 (1-493)

QY 110 GAAGAACGGAGCGATTCATTCAATCTCCAAACCTGCTCTCTGGAC----- 163
Db 104 AspAsnLysGlnValGlnLysHisHisAspLeuValArgMetLeuLeuMetAspGlnAsp 123
QY 164 -----AACATCTCGCAATCGGCACACCCCTCATATAACAACGGGCTATCCGAC 211
Db 124 GlyPheLeuThrGluAsnAsnLysValAspHisPheAspGlyAsnAspLeuTyrAsp 143
QY 212 ATCTACCTGCTCGACGACCCCGACGAAGCCTTGGCGCGCGCCCTTATCGAATCT 271
Db 144 GlnValLeuLysAsp-----IleLysAsn 151
QY 272 GCCGAACACAGCTCGATTGCAATATACATTGCGGCAACGACATTTCCGGCAGGCTG 331
Db 152 AlaLysGluTyrIleHisLeuGluTyrThrPheAlaLeuAspGlyLeuGlyLysArg 171
QY 332 CTGTTCAACCTCATGTACCTTCCCGCAGAACGCGGTGCGGTACGCTGCTGTTGGAC 391
Db 172 IleLeuHisAlaLeuGluLysLeuLysGlnGlyLeuValLysIleLeuTyrAsp 191
QY 392 GACAAACAACACGCGGGGTGGACGATCTCTGCTGCGCCCTCGACAGCCATCCCAATATC 451
Db 192 AspValGlySerLys-----AsnVal 198
QY 452 GAAGTGGCGCTTTCACCCCTTCGTCCTACCAATGCGCGGCGACTCGGC----- 502
Db 199 LysMetAlaAsnPheAsnHisPhe-----LysSerLeuGlyGlyGluVal 213
QY 503 -----TACCTGACCGACTTCGCC-----CGCTCAACCGCGCGCTATCGAC 541
Db 214 GluAlaPhePheAlaSerLysLysLeuProLeuLeuAsnPheArgMetAsnAsnArgAsnHis 233
QY 542 ACAAAATCTTTACCGCCGACAAACCGGCGCCACATACCTGCGCGGACGCAATATCGCGAC 601
Db 234 ArgLysIleIleValIleAspGlyGlnLeuGlyTyrValGlyGlyPheAsnIleGlyAsp 253

QY 602 GAATACTTCAAGTCGGTAGGACACACCGTTTTCGCGACCTGGACATCTCGCCACCGCG 661
Db 254 GluTyrLeuGlyLysLeuGlyTyrTrpArgAspThrHisLeuArgIleGlnGly 273
QY 662 AGCGTCGTCGGCGAAGTATCGACGACTTTCACCGCTACTGGCAAGCACTTCCGCCAC 721
Db 274 AspAlaValAspAlaLeuGlnLeuArgPheIleLeuAspTrpAsnSer---GlnAlaHis 292
QY 722 AACGCCACGCGCATCATC-----CGCAGCGGCAACATCGGCAAG 760
Db 293 ArgProGlnPheGluTyrAspValLysTyrPheProLysLysAsnGlyProLeuGly--- 311
QY 761 GGTCTTCAAGCACTCGGATACAAACGACAAATCCAGACACGCGCTCTGGCTACCGC 820
Db 311 ----- 311
QY 821 GAAACCGTCGAACAGTCGCCCTCTACCAAAAAATACAGCGGACGATCGACTGGCAG 880
Db 312 -----AsnSerProIleGlnIleAlaSerGlyProAlaSerAspTrpHis 327
QY 881 AGCGTCCAA-----ACCGCTGATCAGCAGACACCCCTGCAAAAGGACTCGACCGC 931
Db 328 GlnIleGluTyrGlyTyrThrLysMetIleMetSerAla----- 340
QY 932 GACCGCCCAACCGCGGATTGCGGGAGGCTGCAAGACGCGCTCAACACGCCCGAAAAA 991
Db 341 -----LysLys 342
QY 992 AGCGTCTATCGTTTACCCCTATTTCCTCCCTACAAAAATCCGGCACAGACGCTGGCA 1051
Db 343 SerValTyrLeuGlnSerProTyrPheIleProAspAsnSerTyrIleAsnAlaIleLys 362
QY 1052 AAACGTGTGAGGACGCGCATAGCTTACCTCTGCTGACCAACTGCTCAGCGGACCGAC 1111
Db 363 IleAlaAlaLysSerGlyValAspValHisLeuMetIleProCys-----LysProAsp 380
QY 1112 GTTCCGCGCTCCATTCCGGCTACGTCAAAATACCGAAAAACCGCTCTCAAGCGGCGATC 1171
Db 381 HisProLeuValTyrTrpAlaThrPheSerAsnAlaSerAspLeuLeuSerGlyVal 400
QY 1172 AAATCTTACGAGCTGCAACCCCAACCATCGCTCCCGCCACAAAAAGCAAGGCTGACC 1231
Db 401 LysIleTyr-----ThrTyrGluAsnGlyPhe--- 409
QY 1232 GCGAGCTCCGTAAACGAGCTCGATCGCAAAACCTTCATTGTGCGGCAAGCATCTTC 1291
Db 410 -----IleHisSerLysMetCysLeuIleAspAspGluIleValSer 423
QY 1292 ATCGGCTCATTCACCTCGACCCCGCTTCCGACGCTCAATACCGAAATGGCGTCGTC 1351
Db 424 ValGlyThrAlaAsnMetAspPheArgSerPheGluLeuAsnPheGluValAsnAlaPhe 443
QY 1352 ATCGAAAGCCCAAAATCGCAGAACAGATGGAGCGCACCCCTCGCGATACCAACCCGAA 1411
Db 444 ValTyrAspGluAsnLeuAlaLysAspLeu-----Arg 454
QY 1412 TACGCTTACCGGTTTACCTCGCAAAACACACACCGCTGCAATGGCAGATCCCGCCACC 1471
Db 455 ValAlaTyrGluHisAspIleThrLysSerLysGlnLeu-----Thr 468
QY 1472 CGAAAAACCTACCGAACGAAACCGGACGCAAACTTTTGGAAACGCATCGCGCAAAATC 1531
Db 469 LysGluSerTyrAlaAsnArgProLeuSerValLysPheLysGluSerLeuAlaLysLeu 488
QY 1532 CTATCC 1537
Db 489 ValSer 490

RESULT 30
US-10-429-094-18
; Sequence 18, Application US/10429094
; Publication No. US20030180821A1
; GENERAL INFORMATION:

; APPLICANT: TOMICH, Che-Shen
 ; APPLICANT: QUINN, Cheryl
 ; APPLICANT: ARVIDSON, Staffan
 ; APPLICANT: HARRIS, Douglas
 ; APPLICANT: MOTT, John
 ; TITLE OF INVENTION: ANTIMICROBIAL METHODS AND MATERIALS
 ; FILE REFERENCE: 00774.US1.CN1
 ; CURRENT APPLICATION NUMBER: US/10/429,094
 ; PRIOR FILING DATE: 2003-05-02
 ; PRIOR APPLICATION NUMBER: US 09/966,521
 ; PRIOR FILING DATE: 2001-09-28
 ; NUMBER OF SEQ ID NOS: 145
 ; SOFTWARE: Patent in version 3.0
 ; SEQ ID NO 18
 ; LENGTH: 493
 ; TYPE: PRT
 ; ORGANISM: Staphylococcus aureus
 ; US-10-429-094-18

Alignment Scores:

Pred. No.: 9.27e-11 Length: 493
 Score: 286.50 Matches: 106
 Percent Similarity: 38.2% Conservative: 86
 Best Local Similarity: 21.1% Mismatches: 169
 Query Match: 10.0% Indels: 141
 DB: 4 Gaps: 16

US-10-665-990A-13 (1-1561) x US-10-429-094-18 (1-493)

```

QY 110 GAAGACGGAGGAAAGCGTCAATTAATCTTCAAAACGTCGTCCTCGGAC----- 163
DB 104 AspAsnLysGlnValGlnLysHisAspLeuValArgMetLeuLeuMetAspGlnAsp 123
QY 164 -----AACATCTCGAATCCGGCACACCCCTCATACACAGGGCTATCCGAC 211
DB 124 GlyPheLeuThrGluAsnAsnLysValAspHisPheIleAspGlyAsnAspLeuTyrAsp 143
QY 212 ATCTACCTGCTCGACGACCCCGGAGAGCCCTTCGCGCGCGCCCTTATCGAATCT 271
DB 144 GlnValLeuLysAsp-----IleLysAsn 151
QY 272 GCGAACACAGCCTCGATTGTCATACTACATTTGGCGCAACGACATTTCCGCGAGCGTG 331
DB 152 AlalysGluTyrIleHisLeuGluTyrThrPheAlaLeuAspGlyLeuGlyLysArg 171
QY 332 CTGTTCAACCTCATGTACCTTGGCGCAGACGGCGGTGGCTGACCGCTGCTGTGGAC 391
DB 172 IleuHisAlaLeuGluLysLeuLysGlnGlyLeuGluValLysIleLeuTyrAsp 191
QY 392 GACAAACACACGCGCGGTGGACGATCTCTGTCGCGCTCGACGCCATCCCAATATC 451
DB 192 AspValGlySerLys-----AsnVal 198
QY 452 GAAGTGGCGCTGTTCAACCCCTTGTCTAGCAAAATGGCGCGCACTCGC----- 502
DB 199 LysMetAlaAsnPheAspHisPhe-----LysSerLeuGlyGlyGluVal 213
QY 503 -----TACCTGACCGACTTCCCG-----CGCTCAACCGCGCATGCAC 541
DB 214 GluAlaPhePheAlaSerLysLeuProLeuLeuAsnPheArgMetAsnArgAsnHis 233
QY 542 AACAAATCTTTACCGCGCAACACCGCGCACCATACTCGCGGACGCAATATCCGGAC 601
DB 234 ArgLysIleIleValIleAspGlyGlnLeuGlyTyrValGlyGlyPheAsnIleGlyAsp 253
QY 602 GATATCTTAAAGTCGTGTAGGACACCGTTTTTCGCGCACTGGACATCTCGCCACCGCG 661
DB 254 GluTyrLeuGlyLeuGlyLysLeuGlyTyrTrpArgAspThrHisLeuArgIleGlnGly 273
QY 662 AGCGTCTGGCGGAAGTATCGCAGCACTCGACCGCTACTGGGCAAGCCATTCCGCCAC 721
DB 274 AspAlaValAspAlaLeuGlnLeuArgPheIleLeuAspTrpAsnSer-----GlnAlaHis 292

```

722 AACGCCACGGCGCATCATC-----CGCAGCGCAACATCGGCAAG 760
 293 ArgProGlnPheGluTyrAspValLysTyrPheProLysLysAsnGlyProLeuGly--- 311
 761 GGTCTTCAAGCACTCGGATACAAACGAAACATCCAGACACGCGCTCTCTCGCTACCGC 820
 311 ----- 311
 821 GAAACCGTCAACAGTCGCCCTCTACCAAAAAATACAGAGGGACGCATCGACTGGCAG 880
 312 -----AsnSerProIleGlnIleAlaIleSerGlyProAlaSerAspTrpHis 327
 881 AGCGTCCAA-----ACCGCCTGATCAGCAGACACCCCTGCAAAAGGATCGACCGC 931
 328 GlnIleGluTyrGlyTyrThrLysMetIleMetSerAla----- 340
 932 GACCGCGCAACCGCGGATTGCGGGAGGCTGCAAGACGCGTCAAAACAGCCGCAAAAA 991
 341 -----LysLys 342
 992 AGCGTCTATCTGTTTCAACCTATTTCGTCCTTACAAAATCCGCGCACAGGCACTGGCA 1051
 343 SerValTyrLeuGlnSerProTyrPheIleProAspAsnSerTyrIleAsnAlaIleLys 362
 1052 AAACGTGTGACGACGCGCATAGAGTTCACGCTCTGACCAACTCGCTACAGGCGACCGAC 1111
 363 IleAlaAlaLysSerGlyValAspValHisLeuMetIleProCys-----LysProAsp 380
 1112 GTTGGCGCGCTCATTCGGGTACGTCAATACCGAAACCGCTGTCTCAAGCGCGCATC 1171
 381 HisProLeuValTyrTrpAlaThrPheSerAsnAlaSerAspLeuLeuSerSerGlyVal 400
 1172 AAACCTCTACGAGTGCACCCCAACCATCGCGTCCCGCCCAACAAAGACCAAGCGCTGACC 1231
 401 LysIleTyr-----ThrTyrGluAsnGlyPhe--- 409
 1232 GGCAGCTCCGTAAACCGCTGCATGCAAAACCTTCATTGTGACGCGCAACGCAATCTTC 1291
 410 -----IleHisSerLysMetCysLeuIleAspGluIleValSer 423
 1292 ATCGCTCATATTCACTCGACCCCGTTCGCGACGCTCAATACCGAAATGGCGTGCCT 1351
 424 ValGlyThrAlaAsnMetAspPheArgSerPheGluLeuAsnPheGluValAsnAlaPhe 443
 1352 ATCGAAAGCCCCAAATCCGACAGACAGATGGAGCGCACCCCTCGCCGATACCAACCCGAA 1411
 444 ValTyrAspGluAsnLeuAlaLysAspLeu-----Arg 454
 1412 TAGCGCTACCGCTTACCTCGACAAACACACACCGCTGCAATGGCAGCATCCGCGCAC 1471
 455 ValAlaTyrGluHisAspIleThrLysSerLysGlnLeu-----Thr 468
 1472 CGAAAAACCTACCGCAACGAAACCGGAAACCAACTTTGGAAACGCAATCGCCGCAAAATC 1531
 469 LysGluSerTyrAlaAsnArgProLeuSerValLysPheLysGluSerLeuAlaLysLeu 488
 1532 CTATCC 1537
 489 ValSer 490

RESULT 31

US-09-828-523A-90
 ; Sequence 90. Application US/09828523A
 ; Patent No. US20020168697A1
 ; GENERAL INFORMATION:
 ; APPLICANT: The Pharmacia & Upjohn Company
 ; TITLE OF INVENTION: ANTIMICROBIAL METHODS AND MATERIALS
 ; FILE REFERENCE: 268.62120101
 ; CURRENT APPLICATION NUMBER: US/09/828,523A
 ; CURRENT FILING DATE: 2001-04-06
 ; PRIOR APPLICATION NUMBER: 60/266,327
 ; PRIOR FILING DATE: 2000-04-06
 ; NUMBER OF SEQ ID NOS: 99

OTHER INFORMATION: expression in E. coli
US-09-966-521-84

Alignment Scores:
Pred. No.: 9.3e-11 Length: 502
Score: 286.50 Matches: 106
Percent Similarity: 38.2% Conservative: 86
Best Local Similarity: 21.1% Mismatches: 169
Query Match: 10.0% Indels: 141
DB: 3 Gaps: 16

US-10-665-990A-13 (1-1561) x US-09-966-521-84 (1-502)

```

QY 110 GAAGAACGCGAAGCGTCAATTCATTAATCTCCAAACCTGCTCCTCTGGAC----- 163
   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Db 105 AspAsnLysGlnValGlnLysHisHisAspLeuValArgMetLeuLeuMetAspGlnAsp 124
   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
QY 164 -----ACATCTCGAATCCGGCGACACCCCTCTATAACAGGGCTATCCGAC 211
   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Db 125 GlyPheLeuThrGluAsnAsnLysValAspHisPheIleAspGlyAsnAspLeuTyrAsp 144
   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
QY 212 ATCTACCTGCTCGACGACCCACGAGCCCTTGGCGCCCGCGCCCTTATCGAATCT 271
   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Db 145 GlnValLeuLysAsp-----IleLysAsn 152
   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
QY 272 GCCGAACACAGCCTCGATTTCGAATACATACATTTGGCGCAACGACATTTCCGGCAGGCTG 331
   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Db 153 AlalysGluTyrIleHisLeuGluTyrThrPheAlaLeuAspGlyLeuGlyLysArg 172
   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
QY 332 CTGTTCAACCTCATGTACTTTCGCGCAGAACCGCGGTGCGGTACGCTCTGTGGAC 391
   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Db 173 IleLeuHisAlaLeuGluGluLysLeuLysGlnGlyLeuGluValLysIleLeuTyrAsp 192
   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
QY 392 GACAAACACGCGCGGTGGACGATCTCTGTCTGCGCCCTCGACAGCCATCCCAATATC 451
   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Db 193 AspValGlySerLys-----AsnVal 199
   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
QY 452 GAAGTGCCTGTTCAACCCCTTCGTCTAGCGAAATGGCGCAGCTCGGC----- 502
   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Db 200 LysMetAlaAsnPheAspHisPhe-----LysSerLeuGlyGlyGluVal 214
   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
QY 503 -----TACTGACCGCATTCCTCC-----CCCTCAACCGCGCATGCAC 541
   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Db 215 GluAlaPhePheAlaSerLysLeuLeuAsnPheArgMetAsnAsnArgAsnHis 234
   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
QY 542 AACAAATCTTTACCGCGCAACCGCGCCACCATCTCGCGGACGCAATATCGGCAC 601
   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Db 235 ArgLysIleIleValIleAspGlyGlnLeuGlyTyrValGlyGlyPheAsnIleGlyAsp 254
   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
QY 602 GAATACTTCAAAGTCGGTGAGCACACCGTTTTCGCGACCTGGACATCTCGCCACCGGC 661
   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Db 255 GluTyrLeuGlyLeuGlyLysLeuGlyTyrTrpArgAspThrHisLeuArgIleGlnGly 274
   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
QY 662 AGCGTGTGGCGAAGTATCGACGACTTCGACCGCTACTGGCAAGCCATTCGCCAC 721
   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Db 275 AspAlaValAspAlaLeuGlnLeuArgPheIleLeuAspTrpAsnSer--GlnAlaHis 293
   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
QY 722 AACCCACGCGCATCATC-----CCGACGGGCAACATCGGCAAG 760
   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Db 294 ArgProGlnPheGluTyrAspValLysTyrPheProLysLysAsnGlyProLeuGly--- 312
   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
QY 761 GGTCTTCAAGCACTCGGATACACGAGAAACATCCAGACGCGCTCTCGCGTACCGC 820
   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Db 312 ----- 312
   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
QY 821 GAAACCGTCGAACAGTCCGCTCTACCAAAAAATACAGCGGACGCATCGATCGCGAC 880
   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Db 313 -----AsnSerProIleGlnIleAlaAlaSerGlyProAlaSerAspTrpHis 328
   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
QY 881 AGCGTCCAA-----ACCGCTGTAGCGACACCCCTGCAAAAGGACTCGACCGC 931
   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Db 329 GlnIleGluTyrGlyTyrThrLysMetIleMetSerAla----- 341

```

```

QY 932 GACCGCGCAAAACCGCGAGGTTCGAGAGCGCTCAAGACGCGCTCAAAACAGCCCGCAAAA 991
   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Db 342 -----LysLys 343
   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
QY 992 AGCGTCTATCTGTTTCAACCTATTCCTCCCTACAAAATCCGGCACAGACGCACTGGCA 1051
   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Db 344 SerValTyrLeuGlnSerProTyrPheIleProAspAsnSerTyrIleAsnAlaIleLys 363
   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
QY 1052 AAACCTGGTCAGGACGCGATAGAGTTCACCGTCTCTGACCAACTCGCTACAGCGCAC 1111
   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Db 364 IleAlaAlaLysSerGlyValAspValHisLeuMetIleProCys-----LysProAsp 381
   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
QY 1112 GTTCCGCGCTCCATTCGGGTACGTCAATAACGAAAAACCGCTCTCAAGACCGCGCATC 1171
   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Db 382 HisProLeuValTyrTyrAlaThrPheSerAsnAlaSerAspLeuLeuSerSerGlyVal 401
   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
QY 1172 AAACCTACGAGCTGCAACCAACCATGCGCTCCCGCCCAAAAGAACAAAGAGCGCTGACC 1231
   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Db 402 LysIleTyr-----ThrTyrGluAsnGlyPhe--- 410
   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
QY 1232 GGCAGCTCCGTAAACGCGCTGCATGCCAAAACCTTCATTGTGGAGCGGCAAAAGCATCTTC 1291
   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Db 411 -----IleHisSerLysMetCysteIleAspGluIleValSer 424
   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
QY 1292 ATCGGCTCATTCACCTCGACCCCGCTTCGCGACGGCTCAATACCGAAATGGCGTCGTC 1351
   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Db 425 ValGlyThrAlaAsnMetAspPheArgSerPheGluLeuAsnPheGluValAsnAlaPhe 444
   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
QY 1352 ATCGAAAGCCCCAAATCGCAGACAGATGGAGCGCACCCCTCGCGCATACACACCCGAA 1411
   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Db 445 ValTyrAspGluAsnLeuAlaLysAspLeu-----Arg 455
   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
QY 1412 TAGCGCTACCGGTACCTCGACAAACACACCGCTCGAATGGCAGCATCCGCCAC 1471
   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Db 456 ValAlaTyrGluHisAspIleThrLysSerLysGlnLeu-----Thr 469
   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
QY 1472 CGAAAAACCTACCGCAACGAAACCGAACCAACTTTTGGAAACGCATCGCGCAAAATC 1531
   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Db 470 LysGluSerTyrAlaAsnArgProLeuSerValLysPheLysGluSerLeuAlaLysLeu 489
   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
QY 1532 CTATCC 1537
   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Db 490 ValSer 491
   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
RESULT 33
US-10-429-094-84
; Sequence 84, Application US/10429094
; Publication No. US20030180821A1
; GENERAL INFORMATION:
; APPLICANT: TOMICH, Che-Shen
; APPLICANT: QUINN, Cheryl
; APPLICANT: ARVIDSON, Staffan
; APPLICANT: HARRIS, Douglas
; APPLICANT: MOTT, John
; TITLE OF INVENTION: ANTIMICROBIAL METHODS AND MATERIALS
; FILE REFERENCE: 00774, US1 CN1
; CURRENT APPLICATION NUMBER: US/10/429,094
; PRIOR APPLICATION NUMBER: US 09/966,521
; PRIOR FILING DATE: 2001-09-28
; NUMBER OF SEQ ID NOS: 145
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 84
; LENGTH: 502
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Amino acid sequence encoded by S. aureus coding region cloned from
; OTHER INFORMATION: expression in E. coli
US-10-429-094-84
Alignment Scores:
Pred. No.: 9.3e-11 Length: 502

```

```
Score: 286.50 Matches: 106
Percent Similarity: 38.2% Conservative: 86
Best Local Similarity: 21.1% Mismatches: 169
Query Match: 10.0% Indels: 141
DB: 4 Gaps: 16

US-10-665-990A-13 (1-1561) x US-10-429-094-84 (1-502)

QY 110 GAAGAACGGACGGAAGCGCTCATTTCAATACCTTCAAAACCTGCTCTCTCTGGAC----- 163
   :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
Db 105 AspAsnLysGlnValGlnLysHisAspLeuValArgMetLeuLeuMetAspGlnAsp 124
   :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
QY 164 -----AACATCTGCAATCCGGCACACCCCTCAATAACAACGGGTATCCGAC 211
   :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
Db 125 GlyPheLeuThrGluAsnAsnLysValAspHisPheIleAspGlyAsnAspLeuTyrAsp 144
   :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
QY 212 ATCTACCTCTCGTACGACCCACGAAAGCCCTTGCGCGCGCGCCCTTATCAATCT 271
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 145 GlnValLeuLysAsp-----IleLysAsn 152
   :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
QY 272 GCGAACACAGCTCGATTGCAATACTACTATTGGCGCAACGACATTTCCGCGAGGTG 331
   ||| :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
Db 153 AlaLysGluTyrIleHisLeuGluTyrThrPheAlaLeuAspGlyLeuGlyLysArg 172
   ||| :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
QY 332 CTGTTCAACCTATGATCTTCCGCGACGACGCGGCGTACGCTGCTGTGGAC 391
   :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
Db 173 IleLeuHisAlaLeuGluLysLeuLysGlnGlyLeuGluValLysIleLeuTyrAsp 192
   :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
QY 392 GACAAACACACGCGGGTGGACGATCTCTGCTCGCCCTCGACAGCCATCCCAATATC 451
   ||| :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
Db 193 AspValGlySerLys-----AsnVal 199
   :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
QY 452 GAAGTGGCGCTTTCAACCCCTTGTCTTACGAAATGGCGGCGACTCGGC----- 502
   :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
Db 200 LysMetAlaAsnPheAspHisPhe-----LysSerLeuGlyGlyGluVal 214
   :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
QY 503 -----TACCTGACCGACTTCCC-----CGCTCAACCGCCGCGATGCAC 541
   :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
Db 215 GluAlaPhePheAlaSerLysLeuProLeuLeuAsnPheArgMetAsnAsnArgAsnHis 234
   :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
QY 542 AACAAATCTTTACCGCCGCAACCGCGCCACCATCTCGGCGGAGCAATATCGCGCAC 601
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 235 ArgLysIleIleValIleAspGlyGlnLeuGlyTyrValGlyPheAsnIleGlyAsp 254
   ||| :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
QY 602 GAATACTTCAAGTGGTGGAGACACCGTTTTCGCGGACCTGGACATCTCCGCCACCGC 661
   ||| :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
Db 255 GluTyrLeuGlyLeuGlyLysLeuGlyTyrTrpArgAspThrHisLeuArgIleGlnGly 274
   ||| :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
QY 662 AGCGTGTGCGGGAAGTATCGCAGACTTCCGCGCTACTGGCAGCCATTCGCCGCCAC 721
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 275 AspAlaValAspAlaLeuGlnLeuArgPheIleLeuAspTrpAsnSer---GlnAlaHis 293
   :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
QY 722 AACGCCACGCGCATCATC-----CGCAGCGGCAACATCGGCAAG 760
   :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
Db 294 ArgProGlnPheGluTyrAspValLysTyrPheProLysLysAsnGlyProLeuGly--- 312
   :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
QY 761 GGTCTTCAAGCACTCGGATACACGACGAAACATCCAGACACGCGCTCTCTGCGCTACCGC 820
   :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
Db 312 ----- 312
QY 821 GAAACCGTCGACAGTCCGCTCTTACCAAAAATAACAGCGGACGCTCGACTGGCAG 880
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 313 -----AsnSerProIleGlnIleAlaAlaSerGlyProAlaSerAspTrpHis 328
   :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
QY 881 AGCGTCCAA-----ACCGCGCTGTATCGGACGACACCCCTGCAAAAGACTCGCACC 931
   :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
Db 329 GlnIleGluTyrGlyTyrThrLysMetIleMetSerAla----- 341
   :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
QY 932 GACCGCGCAACCGCGGATTCGCGGAGGCTGCAAGCGGCTCAACAGCGCCGCAAAA 991
   :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
Db 342 -----LysLys 343
   :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
QY 992 AGCGTCTATCTGTTTCAACCTATTTTCTGCTCCCTTACAAAATCCGGCACAGCAGCTGGCA 1051
```

```
Db 344 SerValTyrLeuGlnSerProTyrPheIleProAspAsnSerTyrIleAsnAlaIleLys 363
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1052 AAATGGTGCAGGACGCGATAGCTTACCTCTCTCTGCAACAACTCGCTACAGCGCAGCCGAC 1111
   :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
Db 364 IleAlaAlaLysSerGlyValAspValHisLeuMetIleProCys-----LysProAsp 381
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1112 GTTGGCGCGTCCATTTCGGGTACGTCAAATACCGAAAAACCGTCTCTCAAGCCGCGATC 1171
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 382 HisProLeuValTyrTrpAlaThrPheSerAsnAlaSerAspLeuLeuSerGlyVal 401
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1172 AAATCTACGAGTCAACCCCAACCATCGCTCCCGCCCAACAAAAACAAAGCGCTGACC 1231
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 402 LysIleTyr-----ThrTyrGluAsnGlyPhe--- 410
   :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
QY 1232 GCGAGCTCCGTAACCGAGCTCGATCCGCAAAACCTTCATTGTGGAGCGGCAACGATCTTC 1291
   :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
Db 411 -----IleHisSerLysMetCysLeuIleAspAspGluIleValSer 424
   :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
QY 1292 ATCGCTCATTCACCTCGACCCCGTTCCGACGCGCTCAATACCGAAATGGCGCTGCTC 1351
   :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
Db 425 ValGlyThrAlaAsnMetAspPheArgSerPheGluLeuAsnPheGluValAsnAlaPhe 444
   :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
QY 1352 ATCGAAAGCCCAAAATCGCAGACAGATGAGCGCACCTCGCGCATACACACCCGAA 1411
   :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
Db 445 ValTyrAspGluAsnLeuAlaLysAspLeu-----Arg 455
   :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
QY 1412 TACGCTACCGGTTTACCTCGACAAACACAAACCGCTGCAATGCGACGATCCCGCCACC 1471
   :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
Db 456 ValAlaTyrGluHisAspIleThrLysSerLysGlnLeu-----Thr 469
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1472 CGAAAAAATCTACCCGCAAGCAACCGCCGCAACAACTTTGGAACGCGATCGCGCAAAAATC 1531
   :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
Db 470 LysGluSerTyrAlaAsnArgProLeuSerValLysPheLysGluSerLeuAlaLysLeu 489
   :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
QY 1532 CTATCC 1537
   :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
Db 490 ValSer 491
   :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::

RESULT 34
US-10-084-846A-3
; Sequence 3, Application US/10084846A
; Publication No. US20040006026A1
; GENERAL INFORMATION:
; APPLICANT: WEITNAUER, GABRIELE
; APPLICANT: MUHLNWEIG, AGNES
; APPLICANT: TREFFER, AXEL
; APPLICANT: BECHTHOLD, ANDREAS
; TITLE OF INVENTION: AVILAMYCIN DERIVATIVES
; FILE REFERENCE: 1974-005
; CURRENT APPLICATION NUMBER: US/10/084,846A
; CURRENT FILING DATE: 2003-02-25
; PRIOR APPLICATION NUMBER: PCT/EP01/09815
; PRIOR FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: DE 101 09 166.4
; PRIOR FILING DATE: 2001-02-25
; NUMBER OF SEQ ID NOS: 120
; SOFTWARE: Patentin Ver. 3.2
; SEQ ID NO 3
; LENGTH: 19695
; TYPE: PRT
; ORGANISM: Streptomyces viridochromogenes
; FEATURE:
; OTHER INFORMATION: Protein 1: amino acid sequence encoded by coding strand 1.
; OTHER INFORMATION: Start codon: gga, Start position: nucleotide 1.
US-10-084-846A-3

Alignment Scores:
Pred. No.: 2,92e-10 Length: 19695
Score: 283.00 Matches: 175
Percent Similarity: 33.4% Conservative: 60
Best Local Similarity: 24.9% Mismatches: 226
Query Match: 9.9% Indels: 242
```

DB: 4 Gaps: 36
US-10-665-990A-13 (1-1561) x US-10-084-846A-3 (1-19695)
QY 80 TGTTCATGTTCTTCA---TGGTTG----- 100
DB 5065 CysSerCysSerSerAlaTrpSerSerValThrProThrMetLysAlaArgAlaAlaPro 5084
QY 101 ---CCCCACTGGAAGACGGAGCGAAGCGCTGATTTCAATACCTTCCAAACCTGCTCCTC 157
DB 5085 GlyProPro-TripGluSerSerAlaSerAlaCysArgSer-----ProSerAlaValLe 5102
QY 158 CTGGACAACATCTCAAAATCGGCACACCCCTCATATAACACGGGC-----TATCC 208
DB 5102 uTrpAspThrArgCysThrProAlaThr:ProArgThrAlaSerAlaSerCysArgSerPr 5122
QY 209 GACATCTACCTGCTCGAGCAGCCCGACAGAGCCCTTGGCGCGCGCGCGCTTATCGAA 268
DB 5122 oSerSerTrpAlaSer-----ProAlaProProSerProCysTrpArgL 5137
QY 269 TCTGCCGAACACAGCCTCGATTTGCAATACTACATTTTGGCGCAACGACA----- 317
DB 5137 yCysTrpAsnAlaGly-----CysSerSerProGlyArgAlaProThrProSerTrpGL 5155
QY 318 -----TTTCCGGCAGGCTGCTGTTCAACCTCATGTACCTTTGCCGCGAGAA 361
DB 5155 yProProProTrilleSerProAlaGlyCysCysSerArgValSerSerProCysCysAl 5175
QY 362 CGCGCGTGGCGTACGCTGCTGTTGGACAGACA-----ACAACACGCGGGG 409
DB 5175 aAlaValArgProTrpAlaCysCysAlaProSerArgSerProArgSerSerAlaSe 5195
QY 410 TTGGACGATCTCTCTGCTCGCCC-----TCGACAGCCAT 442
DB 5195 rTrpTrpCysSerSerAlaProProCysAspAlaSerTrpSerGlyArgSerAlaGlyCy 5215
QY 443 CCCAATATCGAAGTGGCGCTGTTCAACCCCTTGGCTTACGCCAAATGGCGCGCATCGGC 502
DB 5215 sProThrAlaAlaCysSerHisSerSerProAlaSerCysCys----- 5229
QY 503 TACCTGACCGACTTCCCGCGCTCAACCGCGCATGCACACAAT----- 548
DB 5230 -----ProArgSerProProSerSerAlaSerIleSerSerSerAlaProSerCysSe 5247
QY 549 -----CCTTTACCGCCG-----AC 562
DB 5247 rAlaProSerAlaArgArgProArgArgProTrpArgThrProAlaAlaSerCysArgAs 5267
QY 563 AACCGCGCCACATAC----- 578
DB 5267 nSerSerProProTyCysCysArgSerSerProArgSerAlaArgProThrCysCy 5287
QY 578 ----- 578
DB 5287 sAsnSerAlaGlyAlaAlaProCysGlyCysGlyProAlaSerProCysTrpTrpPr 5307
QY 579 -----TCGGC 583
DB 5307 oSerSerAlaAsnTrpArgAlaAlaArgArgGlnArgArgCysArgTrpSerAlaSerTh 5327
QY 584 GGACGCAATATCGCGCAGCAATACT-----TC 610
DB 5327 rProCysAlaSerAlaCysSerThrAlaAlaAspProAsnTrpSerSerProSerGlySe 5347
QY 611 AAACCTCGGTGAGGACACG---TTTTCCCGGACCTGGACATCTTCGCCACCGGCA---GC 664
DB 5347 rAsnSerAlaSerHisProArgCysSerProCysTrpSerSerSerProCysAlaProPr 5367
QY 665 GTCCGCGCGCAAGTATCGCAGCTCTCGACCGCTACTGGGCAAGCC---ATTCCGCCAC 721
DB 5367 oSerProArgArgCysSerThrCysSerThrAlaProArgArgAlaProArgProProVa 5387
QY 722 AACGCCACGC-----GCATCATCGCAGCGGCAACATCGGCA----- 758

DB 5387 lGlnLysArgValProSerPheAlaAspAspGlyAlaAlaAsnSerAlaArgGluMetPr 5407
QY 759 ---AGGTCTTCAAGCACTCGATACACACGACGAAACATCCAGACACGCGCTCTCGCGC 814
DB 5407 oGluArgValValIleProSerArgIleProThrLysArgLeuSerSerLysGluCysAr 5427
QY 815 TACCGCAAAACCGTCGACACGTCGCCCC----- 842
DB 5427 gSerValArgAlaGlySerSerLysProGluAsnArgGluValGlyTrpCysArgProTh 5447
QY 843 -----TCTACCAAAAAATACAGACGGGACGACGATCGACTGGC----- 878
DB 5447 rCysThrSerValLeuSerAlaSerSerTyArgArgValAlaSerAsnGlyProTrpAs 5467
QY 879 -----AGAGCGTCCAAACCGCTGTATCAGCAGCACCCCTTGCAAAAGGACTC 925
DB 5467 nAlaAlaSerIleArgProLysArgProArgHisThrSerSerAlaGlySerProProAl 5487
QY 926 GACCGCAGCGCCGCAAAACCGCGATTTGCCGGAGGCTGCAAGACGCGCTCAACAGCGCC 985
DB 5487 aThrSer-ProLeuArgLysTrpProCysGlyArgArgSerArgArgGlyGlyAlaA 5507
QY 986 GAAAAAGCGTCTATCT-----GGTTTCACCC 1012
DB 5507 spArgArgArgAsnSerIleCysCysCysMetArgAlaProGlyThrArgAlaArgThra 5527
QY 1013 TATTTCTGCTTACAAAATCGGCACAGACGACATCGGCAAACTGGTGACAGGCGCAT 1072
DB 5527 laGlyArgPro-----ArgHisThrCysSerGlyThrTrpTrpAlaGlyThrCysS 5544
QY 1073 GACGTTACCGTCTGACCACTCGCTACAGC-----GACC 1108
DB 5544 erProTrpArgSerGlyArgAlaAlaThrAlaCysSerAlaArgTrpAsnSerProSerA 5564
QY 1109 GACGTTCCCGCGTCCATTCGG-----CTACGTCATAATACCGAAACCGCTGTC 1159
DB 5564 laThrCys---ArgProIleArgThrAlaProAlaArgSerSerProProThrThrThrg 5583
QY 1160 AAAGCCCG-----CATCAACTCTACGA----- 1182
DB 5583 luArgArgSerThrGlyGlyGlyTrpAlaProAlaSerSerAlaAlaMetArgSerArgp 5603
QY 1183 -----GCTGCAACCCCAACCATGCGTCCCGCCCAACAAAGACAAAGCGCTGACCGCAGC 1237
DB 5603 roGlyAlaAlaSerGlyProAlaSerProGlySerAlaArgTrp---ProProArgGlyC 5622
QY 1238 TCGTAAACAGCCTGCATGCCAAACCTT-----CATTTGGACGGCAACGCACTTTC 1291
DB 5622 ysArgArgSerSerArgCysThrAlaValThrSerArgCysSerArgArg--AlaSerPr 5641
QY 1292 ATCGGCT---CATTCAACCTCGACCCCGCTTCGCGACGGCTCAATACGAAATGGGCTC 1348
DB 5641 oGlyAlaValProProThrSerAlaProValSerAlaSerSerSerProProAlaAlaPr 5661
QY 1349 GTCATCG---AAAGCCCCAAAATCGCAGAACAGATGGACGCGACCCCTCG---CCGATACC 1402
DB 5661 oProSerLeuArgTrpProArgPheArgThrIleTrpThrProSerAlaProTrpPr 5681
QY 1403 ACACCCG-----AATACGCTACCGGTTACCCCTCGACAAACACAAACCGCGC-- 1448
DB 5681 oGlyProAlaSerGlyTrpArgThrSerProAlaSerProSerThrIleThrArgSerGl 5701
QY 1449 -----TGCAATGGCAGATCCCGCCACCCCGCAAAACCTACCCG 1486
DB 5701 yArgAlaValLeuAsnGlyGlyLeuGlyGlyAlaAlaValProValHisValGlyIleAr 5721
QY 1487 AACGAAACCCGACGAAACCTTTTGAAACGATCGCGCAAAATCTCTATCCCTGCTGCC 1546
DB 5721 gProGlyAspArgProLeuAlaGlyProAlaThr-----IleCysCyPr 5736
QY 1547 A 1547


```
Db 1391 rProThrIlePro-----ThrThrLysProArgCys---TyrProGlySerThrAs 1407
Qy 1178 TACGAGCTGCAACCAACCATCGCGTCCCGCCA-----CA 1213
Db 1407 pThrSerCysGlnProLysThrTyrSerProProThrSerArgProLeuIleThrThrSe 1427
Qy 1214 AAAGACAAAGCGCTGACCGGAGCTCCGTAACACCGCTGCATGCCAAAACCTTCATTGTG 1273
Db 1427 rLysProArgCysTyrProGlySerProAspProGluCysGlnProAlaThrTyrLeu-- 1446
Qy 1274 GACGGCAACGATCTTCATCGCTCATCACTCGACCCCGCTGCGCAGCGCTCAAT 1333
Db 1447 -----ProProThrThrArgThrThrIleProVal-----ThrLy 1459
Qy 1334 ACCGAATGGCGTGCATCGAAAGCCCAAAATCGCAG--AACAGATGAGCGCAC 1390
Db 1459 sProArgCysTyrProGlySerSerAspProGluCysGlnProGluThrTyrSerProPr 1479
Qy 1391 CTCGCGGATACCAACACCGAATACG-----CCTACCGCTTACCTCGACAAACACAAC 1444
Db 1479 oThrArgProProValThrThrSerLysProAsnCysTyrProGlySerThrAspSe 1499
Qy 1445 GCCTGCAATGGCAGATCCCGCACCCGAAACCTTACCGAAGCAACCGAAGCCAAA 1504
Db 1499 rArgCys-----ProGlnLysProProThrThrLeuGlnProLysCysTy 1514
Qy 1505 CTTTGGAAACGCATCGCCGCAA-----AAATCTATCTCCCTGCTGCCCA 1547
Db 1514 rProGlySerSerAspProGluCysLeuAsnCysTyrProGlySerPro 1530

RESULT 36
US-10-684-422-44
; Sequence 44. Application US/10684422
; Publication No. US20040229233A1
; GENERAL INFORMATION:
; APPLICANT: ABURATANI, Hiroyuki
; APPLICANT: YAMAMOTO, Shogo
; TITLE OF INVENTION: Human housekeeping genes and human tissue-specific genes
; FILE REFERENCE: 113991
; CURRENT APPLICATION NUMBER: US/10/684,422
; CURRENT FILING DATE: 2002-10-15
; PRIOR APPLICATION NUMBER: US 60/418,614
; PRIOR FILING DATE: 2002-10-16
; NUMBER OF SEQ ID NOS: 332
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 44
; LENGTH: 820
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-684-422-44

Alignment Scores:
Pred. No.: 2,55e-10 Length: 820
Score: 280.50 Matches: 134
Percent Similarity: 37.8% Conservative: 59
Best Local Similarity: 26.3% Mismatches: 204
Query Match: 9.8% Indels: 113
DB: 5 Gaps: 23

US-10-665-990A-13 (1-1561) x US-10-684-422-44 (1-820)

Qy 217 CTGCTCGACAGCCCAACCAAGCCCTTGGCGCCGCGCCCTTATCGAATCTGCCGA 276
Db 172 ProArgArgLysSerArgSerProSerProArgArgSerProValArgArg 191
Qy 277 ACACGCTCGATTGCACTACTACATTTGGCGCAACGACATTTCCGGCAGGCTGCTGT 336
Db 192 GluArgLysArg-----SerHisSerArgSer----- 200
Qy 337 CAACCTCATGTACTTCGCGCAGAACCGGG---CGTGGCGTACGCTGCTGTGGACGA 393
Db 201 -----ProArgHisArgThrLysSerArgSerProSerProAlaProGluLys 216
```

```
Qy 394 CAACCAACACGCGGGTTGGACGATCTCTCTCGCCCTCGACAGCCATCCCAATATCGA 453
Db 217 LysGlu-----LysThrProGluLeuProGluProSerValLysValLys 231
Qy 454 AGTGGCGCTGTTCACCCCTTCGTCTACG---CAAAATGGCGCGCAGCTCGGTACTCGTAC 510
Db 232 GluProSerValGlnGluAlaThrSerThrSerAspIleLeuValProLysProGlu 251
Qy 511 CGACTTCCCGCTCAACCGCCGATGCAACAATCTTTTACCCTCCGCAACCGCGC 570
Db 252 ProIleProGluProLysGluProSerProLysLysAsnSerLysLysGluGlnGluLys 271
Qy 571 CACCATACTCGG---CGAGCGCAATATCGGCGCAATATCTTCAAGTCGCTGAGGACAC 627
Db 272 GluLysThrArgProArgSerArgSerLys-----SerArgSerArgThr 288
Qy 628 CGTTTTCCCGACCTGGACATCTCTCGCCACCGCGAGCGT----- 666
Db 289 ArgSerArgSerProSerHisThrArgProArgArgArgHisArgSerAspLysMetTyr 308
Qy 667 -----CGTGGCGAAGTATCGACAGCTTCGACCGCTACTGGGCAAGCCATTCGCGCCA 720
Db 309 SerProArgArgProSerProArgArgProSerProArgArgThrProPro 328
Qy 721 CAACGCCACCGCATCATCCGACGCAACATCGCAAGGGTCTTCAAGCACTCGGATA 780
Db 329 ArgArgMetPro-----ProProArgHisArgSerArgSerProValArg 346
Qy 781 CAACGACGA-----AACATCCAGACA 801
Db 347 ArgArgArgSerSerAlaSerLeuSerGlySerSerSerSerSerSerArgSer 366
Qy 802 CGCGCTCTCGGTACCGCGAAACCGTTCGAACAGTCCGCCCTCTTACCAAAAAATACAGAC 861
Db 367 ArgSerProLysLysProLysArgThrSerSerProProArgLys----- 383
Qy 862 GGAACGATCGACTGGCAGAGCGTCCAAACCCGCTCATCAGCAGAC----- 909
Db 384 ---ThrArgLeuSerProSerAlaSerProArgArgArgHisArgProSerPro 402
Qy 910 -----CCCTGCAAAAGGACTCGACCGCGCGCGCAACCGCGATTCGCGG 957
Db 403 ProAlaThrProProLysThrArgAspSerProThrProGlnGlnSerAsnArgThrArg 422
Qy 958 GAGCTGCAAGACGCGCTCAACAGCCGA----- 987
Db 423 LysSerArgValSerValSerProGlyArgThrSerGlyLysValThrLysHisLysGly 442
Qy 988 ---AAAAAGCGTCTATCTGTGTTTCCACCTATTTCCTCCCTACAAAATC----- 1032
Db 443 ThrGluLysArgGluSerProSerProAlaProLysProArgLysValGluLeuSerGlu 462
Qy 1033 ---CGGCACAGACGCACTGGCAAACTGTGTGCGAGA-----CGCATAGACGT 1077
Db 463 SerGluGluAspLysGlyGlyLysMetAlaAlaAspSerValGlnGlnArgGln 482
Qy 1078 TACCGT-----CCTGACCAACTCGCTACAGCGGACCGACGTTGC 1116
Db 483 TyrArgArgGlnAsnGlnSerSerSerSerSerSerSerSerSerSerSerSerSerSer 502
Qy 1117 CGCGCTCATTCGCGCTACGT---CAAATACCGAAACCGCTGCTCAAGCCGCGCATCAA 1173
Db 503 GluArgProLysArgSerHisValLysAsnGlyGluValGlyArgArgArgHisSer 522
Qy 1174 ACTCTACGACTCAACCCCAACCATGCGCTCCCGCCACCAAAAGACAAAGCCCTGACCGG 1233
Db 523 ProSerArgSerAlaSerProSerProArgLysArgGlnLysGluThrSerProArgGly 542
Qy 1234 CAGCTCGTTAACGAGCTGATGCCAAAACCTTCATTGTGGAGCGGCAACCGCATCTTCAT 1293
Db 543 ArgArgArgArgSerProSer-----ProProThrArgArgArg 556
```



```
Db 344 AlaAla-----SerArgPro---ProArgThrArgValProProAlaArgAla 359
QY 1353 TCGAAGCCCAAAATCCAGAACAGATGGAGCGACCTCGCGATACACACCGCAAT 1412
Db 360 SerProSerArgGlyArgSerArgTrpArgProAlaTrpSerGlyAlaProGly 379
QY 1413 AGCCCTACCGCTTACCTCGACAAACACAAACCGCTGCAATGGACGATCCCGCACCC 1472
Db 380 CysProProCysSerProGlySerIleAlaThrSerPheSerCysThrProProPro 399
QY 1473 GAAAAACCTACCGAAGAACACCGGAAG 1499
Db 400 ProProArgArgArgGlyPheProArg 408

RESULT 38
US-10-437-963-134726
; Sequence 134726, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 134726
; LENGTH: 668
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_36471C.1.pep
US-10-437-963-134726

Alignment Scores:
Pred. No.: 3,91e-10 Length: 668
Score: 277.50 Matches: 145
Percent Similarity: 33.9% Conservative: 57
Best Local Similarity: 24.4% Mismatches: 222
Query Match: 9.7% Indels: 171
DB: 4 Gaps: 30

US-10-665-990A-13 (1-1561) x US-10-437-963-134726 (1-668)
QY 33 CCCCGATGAACAGCAGCGCTCATTTCCCTTTATGCC-----TCC 74
Db 43 ProThrGlyGlnHisHisSerSerProSerArgSerTyrlleThrGlnSerLeuPro 62
QY 75 TTCCTGTCTATGTCTTCATGGTTGCCCCACTGGAAGACGACGGAAGCGGTCTATT 134
Db 63 TyrleulleuLeuLeuLeuGlyTyrvAlaSnTrpSerleuGlyGlyGly----- 80
QY 135 TCAATACTTCAAAACCTGTCTCTCTGGACACATCTCTGCAAACTCCGGCACACCCCTCATA 194
Db 80 ----- 80
QY 195 ACAACGGGCTAT-----CCGACATCTACCTGCTCGAGGACCCCGCCAGAGCCC----- 242
Db 81 -----GlyTyAsnProProSerProSerIleGlyThrSerProThrProThrProGlyGly 98
QY 243 -----TTGCGCGCGCGCGCGCCCTTATCGAATCTG 272
Db 99 GlyGlyGlyTyThrProThrProSerAspThrProProSerProSerAspThrSer 118
QY 273 CCGAACAACAGCCTCGATTGTCATTTGCGCGCAACGACATTTCCGGCAGGCTGC 332
```

```
119 ProSerThr-----ProGlyGlyGly 125
333 TGTTCACCTCA-----TGTAACCTTGGCGGAGAACGGCGGTGGCGGTACGCGTCG 383
126 CysSerSerProThrProCysAspAlaProProSerProSerProSerAspThrSer--- 144
384 TGTGGACGACAAACACGCGCGGTGGAGCATCTCTCTCGCCCTCGACGCCATC 443
145 -----ProThrThrProGlyGlyGlyGlyTyThrProThrProSerAspThrPro 162
444 CCAATATCGAAGTCGCGCTGTTC----- 467
163 ProSerProSerSerAspThrSerProThrThrProGlyGlyGlyGlyTyThrPro 182
468 ACCCTTGTCTCTAGCAAAATGGCGGACCTCGGCTACCTGACGACTTCCCGCGCTCA 527
183 ThrProSerAspAlaProProSerProSerSerAspThrProThrThrProGlyGly 202
528 ACCGCGCATGCACAAACAAATCCT-----TTACCGCGGAC-----ACGCG 569
203 GlyGlyGlyTyThrProThrProSerAspAlaProProSerProSerSerAspThrSer 222
570 CCA-----CCATACTCGCGGAGCAATATCGGCG 599
223 ProThrThrProGlyGlyGlyGlyTyThrProThrProSerAspThrProProSer 242
600 ACGAATACTTCAAAAGTCGGTAGACACCG-----TTTTCCCGACCTGG 644
243 ProSerSerGlySerSerProThrThrProGlyGlyGlyGlyTyThrProThrPro 262
645 ACATCTCGCACCGGAGCGTCTCGCGAAGTATCGACGACTTCGACCGCTACTGGG 704
263 SerAspThrProProSerProSerSerGlySerSerArgThrThrProGlyGlyCysSer 282
705 CAAGCCATTCCGCCCAACGCGCATCTCCGCGGCGGCAACATCGGCAAGGTC 764
283 ThrProThrProCysGlyThrProProAlaProSerSerGlyThr----- 297
765 TTCAGCACTCGGATACAAG-----ACGAACATCCAGACACGCGCTCTCGCGCT 815
298 -----SerProThrThrProGlyGlySerTyThrProThrProSerIleGly 314
816 ACCGCGAAACCGTCGAACAGTCGCCCTCTACCAAAATAACAGACGCGGACGATCGACT 875
315 AspValProProSer-----ProSerSerAspThrSerProThrThrProGlyGly 331
876 GGCAGAGCGTCCAAACCCGCGCTGATCAGCGACACCCCTGCAAAAGGACTCGACCGCACC 935
332 GlySerProSerThrProCys-----AspThrPro-----ProSerProSerSer 346
936 GCGCGAAACCGCGATTCGCGGAGGCTGCAAGCGGCTCAACACGCCCGAAAAAGCG 995
347 GlyThrSerProThrThrProGlyGly-----GlyTyTyThrProThrPro 362
996 TCTATCTGGTTTCAACCTATTTCG-----TCCCTACAAATCCGCGCACAG 1040
363 SerValGlyAspValProProSerProAlaSerGlyThrSerProSerProThrProGlySer 382
1041 ACGCACCTGGCAAAACTGGTCAGGACGATAGACGTTACCGTCTCTGACCAACTCGCTAC 1100
383 GlyGlyTyThrPro-----SerThrProCysSerAlaProProSerProSerSerGly 400
1101 AGCGGACCGAGTTCGG-----CCGTCCATTCCGGTACGTCAAAATACC 1145
401 ThrSerProThrThrProGlyGlyTyThrSerProSerThrPro----- 415
1146 GAAACCGCTGCTCAAAAGCGGATCAAACTCTACGAGCTGCAACCCCAACCATCGCGCTCC 1205
416 -----CysAsnAlaProProSerProSerSerAspThrSerProThrThrProGly 432
1206 -----CCGCCACAAAGACA-----AAG 1223
```

Db 433 GlyGlyAsnTyrProProAlaProThrIleGlyAsnValProProSerProSerSerGly 452
QY 1224 GCCTGACGGAGCTCCGTAACACGCTGCATGCCAAACACCTTCAATTGTGGACGGCAAAAC 1283
Db 453 ThrSerProSerThrProGlyGlyCysSerSerProThrProCysAspAlaPro 472
QY 1284 GCATCTCATCGCTCATTAACCTGACGACCGCCCGTTCGCGAGGGCTCA----- 1331
Db 473 ProSerProSerSer---AspThrSerProThrProGlyGlyGlyTyrTyrProPro 491
QY 1332 ATACCGAAATGGCGCTCGTCAAGAACGCCCAAAATCGCAACAGATGAGCGCACCC 1391
Db 492 ThrProSerIleGlyThrSerProSerThrProGlyThrGlyGlyTyrTyrProPro 511
QY 1392 TCGCGG-----ATACACACCGCAATACGCTTACCGCGTTACCTTCGACA 1436
Db 512 SerProSerThrGlyTyrThrProThrProAspValProIleSerThrProSer--- 530
QY 1437 AACACACCGCTGCAATGGCAGCATCCGCCACCGCAAAACCTTACCGGAACGACCG 1496
Db 531 -----SerProTyrSerProLeuValProThrPro-----ProSerSerThrThrPro 546
QY 1497 AAGCCAAATTTGGAAACGATCGCGCGCAAAATCCTATCCCTGC 1541
Db 547 MetProPheAspProAsnThrAlaPro-----PheProCys 558

RESULT 39

US-10-084-846A-8
; Sequence 8, Application US/10084846A
; Publication No. US2004006026A1
; GENERAL INFORMATION:
; APPLICANT: WEITNAUER, GABRIELE
; APPLICANT: MUHLNEMEG, AGNES
; APPLICANT: TREPZER, AXEL
; APPLICANT: BECHTHOLD, ANDREAS
; TITLE OF INVENTION: AVILAMYCIN DERIVATIVES
; FILE REFERENCE: 1974-005
; CURRENT APPLICATION NUMBER: US/10/084,846A
; CURRENT FILING DATE: 2003-02-25
; PRIOR APPLICATION NUMBER: PCT/EP01/09815
; PRIOR FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: DE 101 09 166.4
; PRIOR FILING DATE: 2001-02-25
; NUMBER OF SEQ ID NOS: 120
; SOFTWARE: PatentIn Ver. 3.2
; SEQ ID NO 8
; LENGTH: 19608
; TYPE: PRT
; ORGANISM: Streptomyces viridochromogenes
; FEATURE:
; OTHER INFORMATION: Protein 3: amino acid sequence encoded by coding strand 1.
; OTHER INFORMATION: Start codon: atc, Start position: nucleotide 3.
US-10-084-846A-8

Alignment Scores:
Pred. No.: 6,83e-10 Length: 19608
Score: 277.50 Matches: 140
Percent Similarity: 38.9% Conservative: 67
Best Local Similarity: 26.3% Mismatches: 222
Query Match: 9.7% Indels: 103
DB: 4 Gaps: 23

US-10-665-990A-13 (1-1561) x US-10-084-846A-8 (1-19608)

QY 104 CCACCTGGAAGAGCGGACGAAAGCGCTCATTTCAATA----- 140
Db 5104 ProLeuGluArgArgAspTyrThrArgArg-SerIleSerGlnLeuProValProAspVa 5123
QY 141 -----CTTCCAAACCTCTCTCTGGACAACATCTCGCAAAATCCGGCACACCCCTCAT 193
Db 5123 lHisPheTyrProGlnCysAlaAlaThrAsnLeuAsnTyrGlnTyrGlyAlaProCys-- 5142
QY 194 ACAACGGGGTATCCGACATCTACTCTGCTCGAGACCCCAACGAGCCCTTGGCGCCGC 253

Db 5143 -----ProLysLysCysSerSerAlaValGlyArgArgAspProArgAr 5157
QY 254 GCCGCCCTTATCGAATCTGCCGAACACAGCCTCGAATTTGCAAT-----AC 298
Db 5157 gProProCysSerAlaArgProAlaCysProSerTyrTrpAsnGlyArgProSerArgAl 5177
QY 299 TACATTTGGCCAAACGACATTTCCGGCAGGCTGTGTTCAACCTCATATGACTTCCCGCA 358
Db 5177 aThrThrSerAlaSerArgSerProProAlaAlaProSerSerThrSer-SerAlaA 5197
QY 359 GAAACGGCGTGGCGCTGCTGCTGTTGACGACCAACACACGCG----- 405
Db 5197 rGTPArgSer---ThrProAlaValThrArgSerArgThrAlaCysCysAlaG 5216
QY 406 -----CGGTTGGACGATCTGCTGCTGCCCTCGACAGCCATCCCAATATCGAAGTG 457
Db 5216 lAlaAlaArgThrGlyArgSerThrGlyArg--ArgSerSerValProThrSer--- 5233
QY 458 CGCTGTTCAACCCCTTCTGCTACGNAATGGCGGCGACTCGGCTACTCTGACCGACTTC 517
Db 5234 -----AlaProGlyArgSerThrAlaThr----- 5241
QY 518 CCGCGCTCAACCGCGCATGACACAAATCCTTTACCGCGCACAAACCGCGCCACCAT 577
Db 5242 -ThrSerThrCysSerArgThrArgSerSerArgAlaProThrSerProArgAlaSe 5261
QY 578 CTCGGCGGACGCAATATCGCGACGAAATCTTCAAGTCGGTGAGGACACCGTTTCGCC 637
Db 5261 rArgSerSerAlaCysSerThrAlaThrAlaArgLeuGlyArgSerGlyTyrThrPr 5281
QY 638 GACCTGGACATCTCGCCACCGGAGCTCTGCGGAGAGTATCGCAGACTTTCGACCGC 697
Db 5281 oThr-----ProValSerTrpSerAlaArgSerSerThrThrSer---Se 5295
QY 698 TACTGGGCAAGCCATTCGCGCACAAACGCCACGCGCATCATCCGAGCGGCAACATCGGC 757
Db 5295 rThrProArgAlaGlyProAlaSerSerProSerThrArgThrAlaAlaArgThrArgSe 5315
QY 758 AAGGCTCTTCAAGCAGCTCGGATACAAACGACGAAACATCCAGACACGCGCTCTGCGCTAC 817
Db 5315 rSerAlaThrSerArgSerGlyAlaThrGlyArg-----AlaAlaArgCysCysPr 5332
QY 818 CGCGAAACCGTGAACAGTCGCCCTCTACCAAAAATACAGACGGGAGCGATCCACTGG 877
Db 5332 oValArgProGlnGlyAla-----SerThrSerSerAlaProArgThrAlaGlyThrGl 5350
QY 878 CAGACGTCCTCAAAACCGCTGATCAGCGACACCCCTGCAAAAGGACTCGACCGCACCGC 937
Db 5350 y-----SerArgCysAlaGlyThrGlyThrAlaSerValSerSerAlaThrArgThr-- 5367
QY 938 CGCAACCGCGATTTCCGGGAGGCTGCAAGACGCGCTCAACACAGCCCGCAAAACGCTC 997
Db 5368 -----GlySerTrpSerAlaAlaGlyAsnThrProThrTrpArgAr 5381
QY 998 TATCTGTTTCACTTATTCCTCTACAAAATCCGCGACAGACGCTGCGCAA----- 1052
Db 5381 gCysTrp---ProArgTrpSerArgAsnProArgProCysAlaThrSerTrpArgThrAl 5400
QY 1053 -----AACTGGTCAGGACGCGCATAGAGCTTACCGTCTCTGACCAAC 1093
Db 5400 aSerThrSerArgAlaTyrTrpSerArgThrSer-----ProThrSerProTh 5417
QY 1094 TCGCTACAGGCGACCG---ACGTTGCCCGCTGCATTCGGGTACGTCAATATCCGAAMA 1150
Db 5417 rAlaSerThrAlaProAlaThrThrSerProAlaThrProProAlaSerSerThr--- 5435
QY 1151 CGCTGCTCAAAAGCGCGCATCA-----AACTCTACGAGCTGCAACCAACCATGCC 1201
Db 5436 -ArgCysCysProProAlaSerThrTrpArgSerThrAlaAlaCysTrpProArgPr 5455
QY 1202 GTCCCGCGCA-----CAAAAGACAAAGGCGCTG 1228

Db 5455 oCysSerProArgThrArgThrSerArgArgArgProGlyArgSerThrSerArgSe 5475
QY 1229 ACCGCAGCTCGTAACAGCCTGCATGCCAAACCTTCATTGTGGACGGCAACGCATC 1298
Db 5475 rThraLThrProThrSerGlyCysSerProTrpCysProAlaCysThrSerAsnArgPr 5495
QY 1289 TTCATCGGCTCATCAACCTCGACCCCGTTCCGCACGGCTCAATPACGAAATGGCGTC 1348
Db 5495 oAlaArgProProThrSerSerTrpProThrArgProArgSerAlaProArgProSerTh 5515
QY 1349 GTCATCGAAGCCCAAAATCGCAGAACAGATGGAGCGCACCTCG--CCGATACCACA 1405
Db 5515 rSer-----ArgSerThraLArgTrpProSerProAsnTrpSerArgAspTr 5531
QY 1406 CCCGAATACGCTACCGGTTACCTCGACAAACACACCGCTGCAATGGCACGATCCC 1465
Db 5531 pProThrSerThrProSerProAlaArgThrArgThrAlaProSerArgThPr 5551
QY 1466 GCCACCGGAAACCTACCCGAAACGACCGGAAGCCAAACTTTGGAAAC-----GCATC 1519
Db 5551 oArgProThrProCysArgArgThrThrProSerAlaSerCysSerArgArgProArgTh 5571
QY 1520 GCCGCAAAATCCTATCCTCTGCTGCCA 1547
Db 5571 rProValTrpArgArgProGlyCysPro 5580

RESULT 40

US-10-335-977-9083
; Sequence 9083, Application US/10335977
; Publication No. US20040052799A1
; GENERAL INFORMATION:
; APPLICANT: DOUGLAS SMITH et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES
; RELATING TO HELICOBACTER PYLORI FOR
; DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 10031
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 28 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: Windows NT 4.0
; SOFTWARE: UNIX
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/335,977
; FILING DATE: 30-Dec-2002
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/993,002
; FILING DATE: 17-DEC-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Mandragouras, Amy E.
; REGISTRATION NUMBER: 36,207
; REFERENCE/DOCKET NUMBER: GTN-018
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 742-4214
; INFORMATION FOR SEQ ID NO: 9083:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 206 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: Helicobacter pylori
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (B) LOCATION 1...206

; SEQUENCE DESCRIPTION: SEQ ID NO: 9083:
US-10-335-977-9083

Alignment Scores:
Pred. No.: 5,12e-10 Length: 206
Score: 274.50 Matches: 64
Percent Similarity: 53.4% Conservative: 45
Best Local Similarity: 31.4% Mismatches: 84
Query Match: 9.6% Indels: 11
DB: 4 Gaps: 5

US-10-665-990A-13 (1-1561) x US-10-335-977-9083 (1-206)

QY 947 CCGATTGCGGGAGGCTGCAAGACGCTCAACAGCCGCAAAAGCGTCTATCTGGTT 1006
Db 6 ProllelylleAlaPheGluLysAlaLeuLysAsnAlaLysGluSerValPheIleAla 25
QY 1007 TCACCTATTTCGTCCTTACAAAATCCGCGCACAGACGCACTGGCAAACTGGTCAGGAC 1066
Db 26 SerSerTyrPheIleProGlyLysLysIleMetLysIlePheLysAsnGlnIleSerLys 45
QY 1067 GGCATAGACTTACGCTCTGACCAACTCGCTACAGGCGACCGACGCTTCCGCGCTCCAT 1126
Db 46 GlyIleGluLeuAsnIleLeuThrAsnSerSerThrAspAlaIleValValTyr 65
QY 1127 TCCGGCTACGTCAAAATACCGAAACCGCTCTCAAGCCGCGCATCAAACTCTACGAGCTG 1186
Db 66 GlyAlaTrpGluArgTyrArgAsnLysLeuValArgMetGlyAlaAsnValTyrGluIle 85
QY 1187 CAACCCAACTGCGCTCCCGCCCAAAAGACAAAGCGCTGACCGGACGCTCCGCTAAC 1246
Db 86 ArgAsnAspPhePheAsnArgGlnIleLysGlyArg-----PheSerThrLysHis 102
QY 1247 AGCTGATGCGCAAAACCTTCATTGTGCGAGCGGCAACGATCTTCATCGGCTCATTTAAC 1306
Db 103 SerLeuHisGlyLysThrIleValPheAspAlaLeuThrLeuLeuGlySerPheAsn 122
QY 1307 CTCGACCCCGTTCCGCGACGCTCAATACCGAAATGGCGTCGTCATCGAAAGCCCAAA 1366
Db 123 IleAspProArgSerAlaTyrIleAsnThrGluSerAlaValLeuPheAspAsnProSer 142
QY 1367 ATCGCAGACAGATGGAGCGCACCCCTCCGCGATACCAACACCCGGAATACGCTACCGCGTT 1426
Db 143 PheAlaLysArgValArgLeuSerLeuLysAspHisAla---GlnGlnSerTrpHisLeu 161
QY 1427 ACCCTCGACAAACACACCGCTGCAATGGACGATCCCGCCACCCGCAAAACCTAC--- 1483
Db 162 ValLeuTyrArgHis---ArgValIleTrpGlu-----AlaThrGluGluGlyIleLeu 178
QY 1484 -----CCGAACGAACCCGAGCGCAAACTTTTGAAGACGATCGCGCGCAAAATCCTA 1534
Db 179 IleHisGluLysAsnSerProAspThrSerPhePheLeuArgLeuLysGluTrpSer 198
QY 1535 TCCTGTGTGCC 1546
Db 199 LysValLeuPro 202

Search completed: May 2, 2006, 06:06:03
Job time : 426 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: May 2, 2006, 05:57:05 ; Search time 6.3 Seconds
(without alignments)
2253.681 Million cell updates/sec

Title: US-10-665-990A-13

Perfect score: 2852

Sequence: 1 caaatacaggaatgcgt.....tgccatcgagggtttatta 1561

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 232119 seqs, 45477862 residues

Total number of hits satisfying chosen parameters: 464238

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:
-MODEL=frame+ n2p.model -DEV=xlp
-Q=abss/ABSSWEB spool/US1066590/runat_01052006_112005_9866/app_query.fasta_1
-DB=published_applications AA New -QFMT=fastan -SUFFIX=rapbn -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=40 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -HOST=abs02p
-USER=US1066590 @CGN 1 1 34 @runat_01052006_112005_9866 -NCPU=6 -ICPU=3
-NO MMAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications AA New:*

1: /SID55/ptodata/2/pubpaa/US08_NEW_PUB.pep1.*
2: /SID55/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
3: /SID55/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
4: /SID55/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
5: /SID55/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
6: /SID55/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
7: /SID55/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
8: /SID55/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
9: /SID55/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
10: /SID55/ptodata/2/pubpaa/US11_NEW_PUB.pep.*
11: /SID55/ptodata/2/pubpaa/US11_NEW_PUB.pep.*
12: /SID55/ptodata/2/pubpaa/US60_NEW_PUB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2681	94.0	525	9 US-10-467-657-5462	Sequence 5462, Ap
2	336.5	11.8	5179	11 US-11-108-172-1068	Sequence 1068, Ap
3	294.5	10.3	504	11 US-11-045-004-2518	Sequence 2518, Ap
4	287	10.1	482	11 US-11-045-004-2698	Sequence 2698, Ap
5	250.5	8.8	1970	9 US-10-821-234-1641	Sequence 1641, Ap

Sequence 14646, A
Sequence 11657, A
Sequence 14, Appl
Sequence 84, Appl
Sequence 74, Appl
Sequence 58, Appl
Sequence 104, Appl
Sequence 44, Appl
Sequence 42, Appl
Sequence 142, Appl
Sequence 50, Appl
Sequence 46, Appl
Sequence 60, Appl
Sequence 48, Appl
Sequence 40, Appl
Sequence 52, Appl
Sequence 2, Appl
Sequence 62, Appl
Sequence 5700, Ap
Sequence 7448, Ap
Sequence 7133, Ap
Sequence 109, App
Sequence 14593, A
Sequence 23618, A
Sequence 111, App
Sequence 19864, A
Sequence 28958, A
Sequence 1000, Ap
Sequence 487, App
Sequence 398, App
Sequence 7814, Ap
Sequence 20771, A
Sequence 10, Appl
Sequence 3385, Ap
Sequence 238, App
Sequence 30, Appl
Sequence 12515, A
Sequence 10838, A
Sequence 20039, A
Sequence 17494, A

ALIGNMENTS

RESULT 1
US-10-467-657-5462
; Sequence 5462, Application US/10467657
; Publication No. US20050260581A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON SpA
; APPLICANT: FONTANA Maria Rita
; APPLICANT: PIZZA Mariagrazia
; APPLICANT: MASIGNANI Vega
; APPLICANT: MONACI Elisabetta
; TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/467,657
; PRIOR FILING DATE: 2003-08-11
; PRIOR APPLICATION NUMBER: GB-0103424.8
; PRIOR FILING DATE: 2001-02-12
; NUMBER OF SEQ ID NOS: 9218
; SOFTWARE: SeqWin99, version 1.04
; SEQ ID NO 5462
; LENGTH: 525
; TYPE: PRT
; ORGANISM: Neisseria gonorrhoeae
US-10-467-657-5462

Alignment Scores:
Pred. No.: 1.79e-170
Score: 2681.00
Length: 525
Matches: 518
Percent Similarity: 99.8%
Conservative: 1
Best Local Similarity: 99.6%
Mismatches: 1


```
Query Match: 94.0% Indels: 0
DB: 9 Gaps: 0
US-10-665-990A-13 (1-1561) x US-10-467-657-5462 (1-525)

QY 2 AAAATACAGGCAATCCGCTCTGAATACTATATCCCGCATGAAACACGAGCGCTCATTTCC 61
DB 6 LysileGlnAlaMetProSerGluThrIleSerProMetLysThrArgSerLeuIleSer 25
QY 62 CTTTATGCTCTCTCTCTGTTTCATGTTCTTTCATGTTGCCCGCCACTGGAAGAACGGACG 121
DB 26 LeuLeuCysLeuLeuLeuCysSerCysSerSerTrpLeuProProLeuGluGluArgThr 45
QY 122 GAAACCGCTCATTTCAATCTTCAAACTGTCCTCTGCGACCAACATCTCTGCAAAATCCGG 181
DB 46 GluSerArgHisPheAsnThrSerLysProValLeuLeuAspAsnIleLeuGlnIleArg 65
QY 182 CACACCCCTATAACAGGGGTATCCGACATCTACTGCTCGACGACCCCGACGAGCC 241
DB 66 HisThrProHisAsnGlyLeuSerAspIleIleLeuLeuAspAsnIleLeuGlnIleArg 85
QY 242 CTTGCCCGCCGCGCCCTTATCGAATCTGCGGAACACAGCGCTCGATTGCAATACTAC 301
DB 86 PheAlaAlaArgAlaAlaLeuIleGluSerAlaGluHisSerLeuAspLeuGlnIleArg 105
QY 302 ATTTGGCGAAGCATTTCCGGAGCGTGTGTTCAACCTCATGTACTTGGCCGACAA 361
DB 106 IleTrpArgAsnAspIleSerGlyArgLeuLeuPheAsnLeuMetLeuAlaAlaGlu 125
QY 362 CGCGCGTGGCGTACGCTGTTGGACGACAAACACGCGGGGTGGACGATCTC 421
DB 126 ArgGlyValArgValArgLeuLeuLeuAspAsnAsnThrArgGlyLeuAspLeu 145
QY 422 CTGCTCGCCCTCGACAGCATCCCAATATCGAAGTGGCGCTGTTCAACCCCTCTCGCTA 481
DB 146 LeuLeuAlaLeuAspSerHisProAsnIleGluValArgLeuPheAsnProPheValLeu 165
QY 482 CGAAATGGCGCGCATCTGGCTACTCGACGACTTCCCGCGCTCAACGCGCGATGAC 541
DB 166 ArgLysTrpArgAlaLeuGlyThrLeuThrAspPheProArgLeuAsnArgArgMethis 185
QY 542 AACAAATCTTTACCGCGACAAACCGCCACCATCTCGCGGACGCAATATCGCGCAC 601
DB 186 AsnLysSerPheThrAlaAspAsnArgAlaThrIleLeuGlyArgAsnIleGlyAsp 205
QY 602 GAATATCTTCAAAGTCGAGACACCGTCTTTCGCGACCTGGACATCTCTCGCCACCGC 661
DB 206 GluTyrPheLysValGlyGluAspThrValPheAlaAspLeuAspIleLeuAlaThrGly 225
QY 662 AGCGTGTGGCGAAGTATCGACACGATTCGACCGCTACTGGGCAAGCATTCGCGCCAC 721
DB 226 SerValValGlyValSerHisAspPheAspArgTyrTrpAlaSerHisSerAlaHis 245
QY 722 AACGCCACGCGCATCATCGCAGCGGCAACATCGCAAGGGTCTTCAAGCACTCGGATAC 781
DB 246 AsnAlaThrArgIleIleArgSerGlyAsnIleGlyLysGlyLeuGlnAlaLeuGlyTyr 265
QY 782 AACGCGAAACATCCAGACACGCGCTCTCGCTACTCCGCGAACCGTTCGAACAGTCGCC 841
DB 266 AsnAspGluThrSerArgHisAlaLeuLeuArgTyrArgGluThrValGluGlnSerPro 285
QY 842 CTTCTACCAAAAATACAGCGGAGCGCATCGACTGCGAGAGGTCCAAACCGCGCTGATC 901
DB 286 LeuTyrGlnLysIleGlnThrGlyArgIleAspTrpGlnSerValGlnThrArgLeuIle 305
QY 902 AGCGACACCCCTGCAAAAGACTCGACCGCGACCGCGCAACCCCGCATTCGCGGAGG 961
DB 306 SerAspSerProAlaLysGlyLeuAspArgAspArgLysProProIleAlaGlyArg 325
QY 962 CTGCAAGACGCGCTCAACAGCCCGGAAAGCGTCTATCTGGTTTCAACCTATTTCGTC 1021
DB 326 LeuGlnAspAlaLeuLysGlnProGluLysSerValTyrLeuValSerProTyrPheVal 345
```

```
QY 1022 CCTACAAATCCGCGACAGACGCACTGCGCAAACTGGTGCAGGCGCATAGACGTTACC 1081
DB 346 ProThrLysSerGlyThrAspAlaLeuAlaLysLeuValGlnAspGlyIleAspValThr 365
QY 1082 GTCTGTGACCAACTCGCTACAGCGCACGACGTTGCGCGCTCCATTCGCGGTACGTCAAA 1141
DB 366 ValLeuThrAsnSerLeuGlnAlaThrAspValAlaAlaValHisSerGlyTyrValLys 385
QY 1142 TACCGAAACCGCTCTCAAAAGCCGGCATCAAACTTACGAGCTGCAACCAACCATGCC 1201
DB 386 TyrArgLysProLeuLeuLysAlaGlyIleLysLeuTyrGluLeuGlnProAsnHisAla 405
QY 1202 GTCCCGCGCAAAAGAGCGCTGACCGCGAGCTCCGTAAACGAGCTCGATGCCAAA 1261
DB 406 ValProAlaThrLysAspLysGlyLeuThrGlySerSerValThrSerLeuHisAlaLys 425
QY 1262 ACCTTCTATTGTGACGCGCAAAACGCATCTTTCATCGGCTCATTCACCTCGACCCCGTTCC 1321
DB 426 ThrPheIleValAspGlyLysArgIlePheIleGlySerPheAsnLeuAspProArgSer 445
QY 1322 GCACGGCTCAATACCGAAATGGGTGCTCATCGAAAGCCCAAAATTCGCGAGAACAGATG 1381
DB 446 AlaArgLeuAsnThrGluMetGlyValIleGluSerProLysIleAlaGluGlnMet 465
QY 1382 GAGCGCACCTCGCGGATACCAACCCGCAATACGCTACCGGTTACCGTTCGACCAACAC 1441
DB 466 GluArgThrLeuAlaAspThrProGluTyrAlaTyrArgValThrLeuAspLysHis 485
QY 1442 AACCGCTTGCATGCGACGATCCCGCCACCGCAAAACCTACCGCAACGAAACCGGAGCC 1501
DB 486 AsnArgLeuGlnTrpHisAspProAlaThrArgLysThrTyrProAsnGluProGluAla 505
QY 1502 AAATTTGGAACGATCGCGCAAAATCTATCTATCCCTGCTGCCCATCGAAGGTTTATA 1561
DB 506 LysLeuTrpLysArgIleAlaAlaLysIleLeuSerLeuLeuProIleGluGlyLeuLeu 525

RESULT 2
US-11-108-172-1068
; Sequence 1068, Application US/11108172
; Publication No. US20050260177A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Lodes, Michael J.
; APPLICANT: Secrist, Heather
; APPLICANT: Benson, Darin R.
; APPLICANT: Meagher, Madeleine Joy
; APPLICANT: Stolk, John A.
; APPLICANT: Wang, Tongtong
; APPLICANT: Jiang, Yugu
; APPLICANT: Smith, Carole L.
; APPLICANT: King, Gordon E.
; APPLICANT: Wang, Aijun
; APPLICANT: Clapper, Jonathan D.
; APPLICANT: Skeiky, Yasir A. W.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedrick Thomas S.
; APPLICANT: Carter, Darick
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND DIAGNOSIS
; FILE REFERENCE: 210121.471C15
; CURRENT APPLICATION NUMBER: US/11108,172
; CURRENT FILING DATE: 2005-04-15
; PRIOR APPLICATION NUMBER: US 10/025,380
; PRIOR FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: US 09/922,217
; PRIOR FILING DATE: 2001-08-03
; PRIOR APPLICATION NUMBER: US 09/833,263
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: US 09/649,811
; PRIOR FILING DATE: 2000-08-28
; PRIOR APPLICATION NUMBER: US 09/609,448
; PRIOR FILING DATE: 2000-06-29
; PRIOR APPLICATION NUMBER: US 09/575,251
```

```
; PRIOR FILING DATE: 2000-05-19
; PRIOR APPLICATION NUMBER: US 09/519,444
; PRIOR FILING DATE: 2000-03-06
; PRIOR APPLICATION NUMBER: US 09/504,629
; PRIOR FILING DATE: 2000-02-15
; PRIOR APPLICATION NUMBER: US 09/480,321
; PRIOR FILING DATE: 2000-01-10
; PRIOR APPLICATION NUMBER: US 09/476,296
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1130
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1068
; LENGTH: 5179
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-108-172-1068

Alignment Scores:
Pred. No.:      8,48e-15      Length:      5179
Score:          336.50      Matches:      141
Percent Similarity: 33.6%      Conservative: 55
Best Local Similarity: 24.1%      Mismatches: 198
Query Match:      11.8%      Indels:      190
DB:                11      Gaps:        22

US-10-665-990A-13 (1-1561) x US-11-108-172-1068 (1-5179)

QY 119 ACGGAAAGCGTCAATTTCAAT-----ACTTCCAAACCTGCTCTCGTGGAAC 166
Db 1248 ThrValGluLysHisPheAsnIleCysSerIleThrArgPro-SerThrLeuThrTh 1267
QY 167 ATCTGCGAAATCCGGCACACCCCTATAACAACGGGTATCCGACATCTACCTGCTCGAC 226
Db 1267 rPheThrThrIleThrLeuProThrThrProThrSerPheThrThrThrThrThrTh 1287
QY 227 GACCCACAGAGCCCTTGGCGCGCGCGCCCTTATCGAATCTGCCAACACAGACGCTC 286
Db 1287 rThrProThr-----SerSerTh 1293
QY 287 GATTGTGAATACTATTTGGCGCAACGACATTTCCGGCAGGCTGCTGTTCAACCTCA-- 344
Db 1293 rValLeuSerThrThr-----ProLysLeuCysCysLeuTrpSerAs 1307
QY 344 ----- 344
Db 1307 pTrpIleAsnGluAspHisProSerSerGlySerAspGlyAspArgGluProPheAs 1327
QY 345 -----TGTACCTTGGCGAGAACGGCGGTGGCGTACGCC----- 380
Db 1327 pGlyValCysGlyAlaProGluAspIleGluCysArgSerValLysAspProHisLeuSe 1347
QY 380 ----- 380
Db 1347 rLeuGluGlnHisGlyGlnLysValGlnCysAspValSerValGlyPheIleCysLysAs 1367
QY 380 ----- 380
Db 1367 nGluAspGlnPheGlyAsnGlyProPheGlyLeuCysTyrAspTyrLysIleArgValAs 1387
QY 381 -TGCTGTGG-----ACGACAACAAC 400
Db 1387 nCysCysTipProMetAspLysCysIleThrThrProSerProProThrThrProSe 1407
QY 401 ACGCGCGGGTGGACGATCTCTGCTCGCTCGAGCCATCCCAATATCGAAGTGGCG 460
Db 1407 rProProProThrThrThrThrThrLeuProProThrThrPro----- 1422
QY 461 CTGTTCAACCCCTTCGTCCACGCAATGGCGGCGACTCGCGTACCTGACCGACTCCCC 520
Db 1423 -----SerProProThrThrThrThrThrProProProThrThrProSerProPr 1441
QY 521 CGCCTCAACCGCGCGATGCAACAATACTCTTTACCGCGGACAAACGGCGCCACCATCTC 580
```

```
Db 1441 oIleThrThr-----ThrThrThrProLeuProThrThrThrProSerPro----- 1456
QY 581 GCGGAGCGCATATATCGGAGAGAACTTCAAGTCGGTGAGGACACACGTTTTCGCCGAC 640
Db 1457 -----ProIleSerThrThrThrProProProProProProProProProTh 1474
QY 641 CTGGACATCTCGCCACCGCAGCGTCTGGCGGAAGTATCGCACGACTTCGAC----- 695
Db 1474 rThrThrProSerProProThrThrThrProSerProProThrThrThrThrThrPr 1494
QY 696 -----GCTACTGGCAAGCCATTCGCCCCACACGCCGCGCATCATCCGCGGCGC 748
Db 1494 oProProThrThrThrProSerProProMetThrThrProIleThrProProAlaSerTh 1514
QY 749 AAGATCGGCAAGGTCTTCAAGCACTCGGATACAGCAACGAAACATCCAGACACGCGTC 808
Db 1514 rThrThrLeu-----ProProThrThrThrProSerProProThrThrTh 1529
QY 809 CTGCGCTACCGCGAAACCGTGAAC-----AGTCGCCCTCTACCAAAAATAACAGC 862
Db 1529 rThrThrThrProProProThrThrThrProSerProProThrThr----- 1544
QY 863 GGACGCATCGACTGGCAGAGCGTCCAAACCCGCGCTGATCAGCGCACACCCCTGCCAAAAGGA 922
Db 1545 -----ThrProIleThrProPr 1550
QY 923 CTGACCGCGACCGCGCAAAACCGCGATTCGCGGAGGCTGCAAGACGCGCTCAACAG 982
Db 1550 oThrSerThrThr-----LeuPro-----ProThrThrThrProSe 1563
QY 983 CCGGAAAAAGGCTCTATCTGGTTTACCCCTATTTCGTCCTACAAATACCGCACAGAC 1042
Db 1563 rProProProThrThrThrThrThrProProProThrThrThrProSerProThrTh 1583
QY 1043 GCATGCGCAAAACTGGTGCGAGGCGCATAGAGCTTACCGCTCTCGTCAAAAT-----AC 1093
Db 1583 r-----ThrThrProSerProProThrIleThrTh 1593
QY 1094 TCCTACAGCGGACGAGCGTGGCGGCTCCATTCGGCTACGTCATAAT----- 1142
Db 1593 rThrThrProProProThrThrThrProSerProProThrThrThrThrThrProPr 1613
QY 1143 -----ACCGAAAAACCGCTGCTCAAGCGCGCATCAAACTC 1177
Db 1613 oProThrThrThrProSerProProThrThrThrProIleThrProProThrThrTh 1633
QY 1178 TAGAGCTGCAACCCCAACCATGCGTCCCGCGCACAAAGACAAAGGCGCTGACCGGCGC 1237
Db 1633 rThrLeuProProThrThrThrProSerProProProThrThr-----ThrThrTh 1650
QY 1238 TCCTTAACCGCTCGATGCGCAAAACCTTCATTGTGGAGCGGCAAGCATCTTCATCGCG 1297
Db 1650 rProProProThrThrThrProSerProProThr-----ThrThrProSerProPr 1669
QY 1298 TCATTCAACCTCGACCCCGTTCGCGACGCGCTCAATACCGAAATGGGCGTCGTCATCGAA 1357
Db 1669 eThrThrThrThrProProProProThrThrThrProSerProIleThrThrThrPr 1689
QY 1358 AGCCCCAAATCCGACAGACAGATGGAGCGCACCTCGCGGATACCAACACCGCAATACGCC 1417
Db 1689 oSerProProThr-----ThrThrMetThrThrProSerProThrThrThrProSerSerPr 1708
QY 1418 TACCGGTTACCTCGACAAAACAAACGCGCTGCAATGGACGATCCCGCCACCGGAAAA 1477
Db 1708 oIle-----ThrThrThrThrThrProSerSerThrThrThrProSerProPr 1725
QY 1478 ACCTACCGCAAGAACCCGAGCAAACTTTGGAACGCGATCGCCGCAAAAATCCTATCC 1537
Db 1725 oThrThrMetThrThrProSerProThrThrThrProSerProProThrThrThrMetTh 1745
QY 1538 CTGCTGCCCA 1547
```



```
QY 746 GGCACATCGCAGAGGCTTCAAGCA-----CTCGGATACAAGCAGAAACATCC 796
|||
Db 263 GlyLysAlaValTyrAlaMetGlnThrArgPheIleMetAspTrpAsnSerAlaSerSer 282
|||
QY 797 AGACACGGCTC-----CTGGCTACCGCGAAACCGTC-----GAACAGTCGCC 841
|||
Db 283 ThrHisLysIleAspTyrLysAlaArgTyrPheProThrPheHisGlyLysGlyHisThr 302
|||
QY 842 CTCTACCAAAAATACAGACGGGA---CGCATCGACTGGCAGAGCGTCCAAACCGCCTG 898
|||
Db 303 SerMetGlnIleValSerSerGlyProAspSerGluTrpGlnGlnIleLysAsnGlyTyr 322
|||
QY 899 ATCAGCGACACCCCTGC AAAAGACTCGACCGGACCGCGCAAAACCCGATTCGCCGG 958
|||
Db 323 IleLys----- 324
QY 959 AGGCTGCAAGACGGCTCAAAACAGCCCGAAAAAGCGTCTATCTGGTTTACCCCTATTTC 1018
|||
Db 325 -----MetIleAsnAlaLysLysThrIleTyrLeuGlnSerProTyrPhe 340
|||
QY 1019 GTCCCTCAAAAATCCGGCAGACGACTGGCAAAAACCTGGTGCAAGACGGCATAGAGTT 1078
|||
Db 341 IleProAspAlaSerLeuLeuGluAlaIleLysIleAlaAlaLeuSerGlyValAspVal 360
|||
QY 1079 ACCGTCTGACCAACTCGCTACAGCGGACCGACGCTTGGCGCGTCCATTCGGGTACGTC 1138
|||
Db 361 ArgValMetIleProAsn-----LysProAspHisAlaPheValTyrArgAlaThrThr 378
|||
QY 1139 AAATACCGGAAAAACCGCTGCTCAAGCGCGCATCAAACTCTACGAGCTCAACCCCAACCAT 1198
|||
Db 379 AsnTyrAlaGlyGluLeuMetGluThrGlyAlaLysIlePhe----- 392
|||
QY 1199 GCCGTCCCGCCACAAAAAGCAGCAAGAGCGCTGACCGGACGCTCCGTAAACAGCGCTGCATGCC 1258
|||
Db 393 -----IleTyrAspAsnGlyPhe-----IleHisAla 401
|||
QY 1259 AAACCTTCATTGTGGAGGCAAGCGCATCTTCATCGGCTCATCAACTCGACCCCGT 1318
|||
Db 402 LysThrLeuValValAspGlyGluIleAlaSerValGlyThrAlaAsnMetAspPheArg 421
|||
QY 1319 TCCGCACGGCTCAATACCGAAATGGCGTGTGTCATCGAAAGCCCAAAATCGCAGAACAG 1378
|||
Db 422 SerPheArgLeuAsnPheGluValAsnAlaPheIleTyrGluLysGlnMetValGlnLys 441
|||
QY 1379 ATGAGCGCACCCCTCGCCGAT-----ACCAACCCGAA---TAC 1414
|||
Db 442 LeuGluAspAlaPheLeuGluAspIleLeuLysSerTyrGlnLeuThrProGluLeuTyr 461
|||
QY 1415 GCCTACCGC 1423
|||
Db 462 AlaLysArg 464
|||
```

RESULT 5

```
US-10-821-234-1641
; Sequence 1641, Application US/10821234
; Publication No. US20050255114A1
; GENERAL INFORMATION:
; APPLICANT: Labat, Ivan
; APPLICANT: Stache-Crain, Birgit
; APPLICANT: Andarmani, Susan
; APPLICANT: Tang, Y. Tom
; TITLE OF INVENTION: Methods for Diagnosis and Treatment of Preeclampsia
; FILE REFERENCE: 821A
; CURRENT APPLICATION NUMBER: US/10/821,234
; CURRENT FILING DATE: 2004-04-07
; PRIOR APPLICATION NUMBER: US 60/462,047
; PRIOR FILING DATE: 2003-04-07
; NUMBER OF SEQ ID NOS: 1704
; SOFTWARE: pt_seq_genes Version 1.0
; SEQ ID NO 1641
; LENGTH: 1970
; TYPE: PRT
; ORGANISM: Homo sapiens
```

```
US-10-821-234-1641
Alignment Scores:
Pred. No.: 3,79e-09 Length: 1970
Score: 250.50 Matches: 123
Percent Similarity: 38.2% Conservative: 69
Best Local Similarity: 24.5% Mismatches: 196
Query Match: 8.8% Indels: 115
DB: 9 Gaps: 24
US-10-665-990A-13 (1-1561) x US-10-821-234-1641 (1-1970)
QY 108 TGGAGAACGAGCAGGAAGCGTCAATTCAATATCTTCAAACTGTCTCTCTGACACA 167
|||
Db 1520 TrpAsnGlnGly-----AlaThrProAlaTyrGlyAlaTrpSerPro 1533
|||
QY 168 TCCTGCAAAATCCGGCACACCCCTCATAAACAGCGGTATCCGACATCTACCTGC----- 221
|||
Db 1534 SerValGlySerGlyMetThrProGlyAlaAlaGlyPheSerProSerAlaSerAsp 1553
|||
QY 222 -----TCGACGACCCCCACGAAGCCCTTG 245
|||
Db 1554 AlaSerGlyPheSerProGlyTyrSerProAlaTrpSerProThrProGlySerProGly 1573
|||
QY 246 CCGCGCGCGCGCTTATCGAATCTGCGAACAACAGCCTCGATTGCAATACATATT 305
|||
Db 1574 SerProGlyProSerSerProTyrIleProSer----- 1584
|||
QY 306 GGCACACGACATTTCCGGCAGCGTGTGTTCAACCTCATATGCTTCCCGCAGAACCG 365
|||
Db 1585 -----ProGlyGlyAlaMetSerProSerTyrSer--ProThrSerPro 1598
|||
QY 366 GCGTGGCGGTACGCTGCTGTGGACGACAACACGCGCGGTGTGACGATCTCCTGC 425
|||
Db 1599 Ala-----TyrGluProArgSerProGlyGlyTyrThrProGln--- 1611
|||
QY 426 TCGCCCTCGACAGCCATCCCAATATCGAAGTGGCGCTGTTCACCCCTTCGTCTACGCA 485
|||
Db 1612 SerProSerTyrSer--ProThrSerProSerTyrSerProThrSerProSerTyrSer 1630
|||
QY 486 -----AATGGCGGCACTCGGCTACCTGACCGACT---TCCCGCGCTCAACCGCCGCA 536
|||
Db 1631 ProThrSerProAsnTyrSerProThrSerProSerTyrSerProThrSerProSerTyr 1650
|||
QY 537 TGCACAAATAATCTTTACCGCCGACAAACCGCGCCACCATCTCGCGGAGCGCAATATCG 596
|||
Db 1651 SerProThrSerProSerTyrSerProThrSerProSerTyrSerProThrSerProSer 1670
|||
QY 597 GCGACGAATACTTCAAAAGTCGGTGAGGACACCGTTTTTCGCGACCTGGACATCTCTGCGCA 656
|||
Db 1671 TyrSerProThrSerProSerTyr--SerProThrSerProSerTyrSerProThrSer 1689
|||
QY 657 CCGCAGACGCTCG-----TCGGCGAAGTATCGCACGACTTCGACGCTACTGGGCAAGCC 710
|||
Db 1690 ProSerTyrSerProThrSerProSerTyrSerProThrSerProSerTyrSerProThr 1709
|||
QY 711 ATTCCGCCCAACGCGCACGCGCATATCCCGCGGCAACATCGCGAAGGGTCTTCAAG 770
|||
Db 1710 SerProSerTyrSerProThrSerProSerTyrSerProThrSerPro-----Ser 1726
|||
QY 771 CACTCGGATACAACGACGAAACATCCAGACACGCGCTCTCGCTACCGCGAAACCGTCG 830
|||
Db 1727 TyrSerProThrSerProSerTyrSerProThrSerProSerTyrSerProThrSerPro 1746
|||
QY 831 AACAGTCCGCCCTCTACCAAAAAATACAGCGGACGCGCATCGACTGGCAGAGCGTCCAAA 890
|||
Db 1747 AsnTyrSerProThrSerProAsnTyrThrProThrSerProSerTyrSerProThrSer 1766
|||
QY 891 CCGCGCTGATCAGCGACACCCCTGCAAAAGACTCGACCGGACCGCGCAACCCGCA 950
|||
Db 1767 Pro-----SerTyrSerProThrSerProAsnTyrThrProThrSerProAsn----- 1782
|||
QY 951 TTGCGGGAGGCTGCAAGACGCGCTCAAAACAGCCGCGCTCAAAAGGCGTCTCTCTGTTTCAC 1010
|||
```



```
QY 1027 AAAATCCGCGCAGACGCTACGCAAAACAACTGGTGGCAGGA----- 1065
Db 324 ArgLeuArgHisAspArgAlaProGlnAlaGlyProGlyHisArgGlyAlaProArgAla 343
QY 1066 -----CGCATAGAGTTACCGTCTGACCAACTCGCTACAGGCGCAGCGTTGC 1116
Db 344 GlnLeuArgArgHisGlnGlnAlaValProGlnValLeuLeuGlnAlaArg----- 361
QY 1117 CGCCGCTCCATTCGGCTACGTCAAATACCGAAACCGCTGCTCAAAGC----- 1164
Db 362 ArgHisGlnValArgArgArgGlyAspProAlaGlyAlaAlaHisAspProArgLeu 381
QY 1165 ---CGGCATCAAACTTACGAGCTGCAACCCCAACCATGCCGTCC----- 1206
Db 382 ArgHisArgGluLeuArgGlyCysAlaArgPro-----ProGlyValGlnAla 399
QY 1207 -----CGCCAAAGACAAAGCGCTGACCGGCGCTC 1239
Db 400 ProGlyAspAlaAlaArgValAlaArgGlyAlaGlyGlnGlyGluAspValAlaVal 419
QY 1240 CGTACACGCTGCTGCAAAAC----- 1263
Db 420 ArgArgAlaProGlnAlaArgArgGlyArgValGlyAspGlyValGlyGlyArgAlaGly 439
QY 1263 ----- 1263
Db 440 LeuLeuLeuProProValGluArgValGlyArgGlyAspGlyArgGlyGlyGly 459
QY 1264 -----CTTCATGTGGCGGCAACGCTCTTCATCGGCTCATTCAA----- 1305
Db 460 AspArgLeuLeuHisAspProArgLeuHisLeuGlnArgValArgAlaProGly 479
QY 1306 -----CCTCGACCC 1314
Db 480 GluArgAlaAspArgAspProProGlyArgAlaHisGlyProValHisAlaProArgArg 499
QY 1315 CGTTCCGCGCAGCTCAATACCGAAATGGCGTGTCTATCGAAAGCCCAAAAT----- 1368
Db 500 ProAlaAlaValAlaAlaGlyGluProGlyGlyGlyHisGlyGluProGlnProAlaGly 519
QY 1368 ----- 1368
Db 520 ProArgGluProValArgValProArgGlyGlyAlaValAlaGlnGlyValGlyLeu 539
QY 1369 CGCAGAACAGATGGAGCGCCTCGCGATACCAACCCGCAATACGCTACCGGTTAC 1428
Db 540 ArgGlnGlyProValHisGlyArgAlaHisGlnValArgValArg-----ArgGly 557
QY 1429 CTTGACAAACACACCGCTGCAATGGCAGCATCCGCGCACCCGAAACCTACCCGAA 1488
Db 558 ProValArgArg-----AlaLeuLeuArgSerHisGlyProGlyArgGlyProPro 575
QY 1489 CGAACCCGAGCAAACTTGGAAACGATCCCGCAAAATCCTATCCCTGCTGCCCAT 1548
Db 576 AlaArgArgGlyArgValArgAlaHisLeuArg-----ProArgArgAla 591
QY 1549 CGAAGCTTATT 1560
Db 592 ArgArgHisVal 595
```

RESULT 7

```
US-11-096-568A-11657
; Sequence 11657, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nikolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
```

```
; SEQ ID NO 11657
; LENGTH: 493
; TYPE: PRT
; ORGANISM: Triticum aestivum
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(493)
; OTHER INFORMATION: Ceres Seq. ID no. 15219800
US-11-096-568A-11657
```

```
Alignment Scores:
Pred. No.: 4,61e-09 Length: 493
Score: 248.00 Matches: 126
Percent Similarity: 35.5% Conservative: 61
Best Local Similarity: 23.9% Mismatches: 163
Query Match: 8.7% Indels: 178
DB: 11 Gaps: 18

US-10-665-990A-13 (1-1561) x US-11-096-568A-11657 (1-493)
QY 34 CCCGATGAAACACGACGCTCATTTCCCTTTTATGCCTCTCTCTGTTCATGTTCTTC 93
Db 28 ProAspProSerAlaSerProProProProProProProProProProProProPro 47
QY 94 ATGGTTGCCCTGGAAGAACGACGGAAG---CCGTATTTCATCTACTTCCAAACC 150
Db 48 AlaSerSerProSerThrArgThrAlaGlyAlaLeuProThrCysArgAlaArgArg 67
QY 151 TGTCTCTCTGGCAACATCCTGCAATCCGGCACACCCCTCATAACAACGGGCTATCCGA 210
Db 68 CysTrpAlaGly----- 71
QY 211 CATCTACTGTCGACGACCCCGAAGCCTTGGCCCGCGCGCCGCTTATCGAATC 270
Db 72 -----ProSerArgArg---ProArgSerProSerArgProSerProAlaProThrSer 88
QY 271 TGCCGACACAGCTCGATTTGCAATACTACTATTGGCGCAACGACATTTCCGGCAGGCT 330
Db 89 TrpAlaAlaSerProThrGlyArgAlaThrArg---AlaThrArgAlaArgThrAlaThr 107
QY 331 GCTGTTCACCTCATGTACTTGGCGCAGAACGCGCGTACGCGTCTGCTGTGA 390
Db 108 ThrAlaArgProProProAlaThrArgSerSerThrSerProArgProThrThr--- 126
QY 391 CGAACAAACACGCGGGTTGGACGATCTCTGCTCGCCTCGACAGCATCCCAATAT 450
Db 127 -----ArgSerProProProProProProProProProProProProProSerAla 138
QY 451 CGAAGTGGCTGTTCACAC--CCTTCGTCCTACGCAATGGCGCGCATCGGCTACCTGA 509
Db 139 SerSerThrProSerArgProSerThrArgAlaSerGlyProSerGlyAlaSer--- 157
QY 510 CGGACTTCCCGCTCAACCGCCGATGCAACAACAAATCCTTTACCGCGCAACCGCG 569
Db 158 ThrSerGlyProSerSerProSerAlaAlaThrArgArgSerArgProSerAlaAlaArg 177
QY 570 CCACCATCTCGGCGAGCAATATCGCGCAGCAAT----- 605
Db 178 ProArgArgSerAlaProAlaSerAlaThrThrArgThrThrAlaProThrThrThr 197
QY 606 -----ACTTCAAAGTCGTGAGGACACCGTTTTCGCGCACCTGGACATCTCGCCACCG 659
Db 198 ArgAlaSerAlaAlaThrSerArgThrProProProProProProProProProProThr 217
QY 660 GCAGCGCTCGTGGCGAAGTATCGCACGACTTTCGACCGCTACTGGGCAAGCCATTCGCGC 719
Db 218 GlyProSerSerArgSerProSerProThrSerProSerProSerProSerProSerPro 237
QY 720 ACAACGCGCAGCGCATCATCGCGACGCGCAACATCGGCAAGGCTCTTCAAGCACTCGGAT 779
Db 238 ThrSerProArgThrCysSerSerAlaAlaProSerThrArg----- 251
QY 780 ACAACGACGAAACATCCAGACACGCGCTCTCGCTACCGGAAACCGTCGAAACAGTCGC 839
```



```
Db 252 -----ThrThrAlaProThrThrAlaSer-----:::ThrProArg 263
QY 840 CCCTCTACCAAAATACAGACGCGACGCTGCGAGCGTCCAAACCCGCTGA 899
Db 264 ProAlaAlaSerSerAlaSerArgThrAlaSerSerArgSerProPro-----281
QY 900 TCAGCGACACCCCTGCAAAAGAGACTCGACCGCGACCGCGCAAAACCGCGATTGCGCGGA 959
Db 282 ---ThrThrPro-----284
QY 960 GGCTGCAAGACGCGCTCAACAGCCGCAAAAGCGTCTATCTGTTTACCCCTATTTCG 1019
Db 285 -----Ser 285
QY 1020 TCCCTACAAATCCCGGCACAGACGCACTGGCAAAACTGGTGCAGGACGCATAGACGTTA 1079
Db 286 SerAlaAlaSerPro-----290
QY 1080 CGTCTGACCAACTCGCTACAGCGACCGACGTTGCGCGCTCCATTCCGCTACGTCA 1139
Db 291 -----ArgArgThrArgProProSerSerProProThrPro 302
QY 1140 AATACCGAAACCGCTGCTCAAGCGCGCATCAAACTCTACGAGCTGCAACCCCAACCATG 1199
Db 303 SerSerProProSerCysAlaArgProAlaAlaSerSerProGly-----317
QY 1200 CGTCCCGCCGACAAAGACAAAGCGCTGACCGGACGCTCGTAACGACGCTGCATGCCA 1259
Db 317 -----317
QY 1260 AAACCTTCATTGTGACGGCAAAAGCACTTCCTCATCGGCTCATCAACCTCGACCCCGGTT 1319
Db 318 -----ThrLeuSerCysSerAlaSerAlaThrSerSerSer-----329
QY 1320 CGGACGCGCTCAATACCGAATGGCGTCTCATCGAAGCGCCCAAAATCG-----1370
Db 330 -----SerThrSerAlaArgAlaProSerSerIleCysSer 341
QY 1371 ---CAGAACAGATGAGCGCACCTCGCGGATACACACCGCAATA---CGCTACCGCG 1424
Db 342 LeuSerThrArgLeuArgSerSerSerLeu---ArgThrArgMetThrSerIleProAr 361
QY 1425 TTACCCTCGACAAACACACACCGCTGCAATGGCAGCATCCCGCCACCGCAAAACCTACC 1484
Db 361 gArgProLeuLeuSerArgProProThrSerThrArgThr-----374
QY 1485 CGAAGGACCGAAGCCAACTTTGG-----AAAGCATCGCCGCAAAATCC 1532
Db 375 -----SerArgSer-ArgCysTrpCysValMetAlaArgArgLeuProLeuMetSerL 392
QY 1533 TATCCCTGCTGCCATC 1549
Db 392 euthrArgLeuProLeu 397

RESULT 8
US-11-169-232-14
; Sequence 14, Application US/11169232
; Publication No. US20060025570A1
; GENERAL INFORMATION:
; APPLICANT: Turner, Katherine
; Clark, Stephen C.
; Jacobs, Kenneth
; Hewick, Rodney M.
; Gesner, Thomas G.
; TITLE OF INVENTION: Megakaryocyte Stimulating Factors
; NUMBER OF SEQUENCES: 143
; CORRESPONDENCE ADDRESS:
; STREET: 87 CambridgePark Drive
; CITY: Cambridge
; STATE: Massachusetts
; COUNTRY: U.S.A.
```

```
ZIP: 02140
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/11/169,232
FILING DATE: 28-Jun-2005
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/10/124,557
FILING DATE: 16-Apr-2002
APPLICATION NUMBER: US 07/643,502
FILING DATE: 18-Jan-1991
APPLICATION NUMBER: US 07/546,114
FILING DATE: 29-JUN-1990
APPLICATION NUMBER: US 07/457,196
FILING DATE: 29-DEC-1989
APPLICATION NUMBER: US 07/390,901
FILING DATE: 08-AUG-1989
ATTORNEY/AGENT INFORMATION:
NAME: Cseri, Luann
REGISTRATION NUMBER: 31,822
REFERENCE/DOCKET NUMBER: GI 5190
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)876-1170
TELEFAX: (617)876-5851
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 941 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 14:
US-11-169-232-14

Alignment Scores:
Pred. No.: 5,42e-09 Length: 941
Score: 247.50 Matches: 135
Percent Similarity: 35.2% Conservative: 44
Best Local Similarity: 26.5% Mismatches: 205
Query Match: 8.7% Indels: 125
DB: 11 Gaps: 29

US-10-665-990A-13 (1-1561) x US-11-169-232-14 (1-941)
QY 144 CCAAACTGCTCTCTCGCAACAATCTGCAAAATCCGGCACACCCCTCATAACAACGGGC 203
Db 132 ProThrProLysAlaGluThrThrThr-----LysGlyProAlaLeuThrThrProLys 149
QY 204 TATCCGACATCTACC-----TGCTGACGAGCCCGCCAGAGCCCTTGCGC 248
Db 150 GluProThrProThrThrProLysGluProAlaSerThrThrProLysGluProThrPro 169
QY 249 -----CCGCGCGCCGCTTATCGAATCTGCGAACAACAGACGCTCGATT 290
Db 170 ThrThrIleLysSerAlaProThrThrProLysGluProAlaProThrThrThrLysSer 189
QY 291 TGCAATACTACA-----TTTGGCGCAACGACATTT 320
Db 190 AlaProThrThrProLysGluProAlaProThrThrThrLysGluProAlaProThrThr 209
QY 321 CCGGACGAGTGTGTTCAACCTCATCTTGGCGGAGACGCGGCTGGCGGTACGCC 380
Db 210 ProLysGluProAlaProThrThrThrLysGluProAla-----222
QY 381 TGCTGTTGGACGACAAACAACACGCGCGGTTTGACGATCTCTGCTCGCCCTCGACAGCC 440
Db 223 -----ProThrThrThrLysSerAlaProThrThrProLysGluProAlaProThr 239
QY 441 ATCCCAATATCGAAGTGGCGCTGTTCACCCCTTGGTCTTACGCAATGGCGGACATCG 500
```

Db 240 ThrPro---LysLysProAlaProThrThrProLysGluProAlaProThrThrProLys 258
QY 501 GCTACTGACCTTCCCGCCCTCAACCGCGCATGCACAAACAAATCTTTACCG--- 557
Db 259 GluProThrThrProLysGluProAlaProThrThrLysGluProAlaProThr 278
QY 558 -----CCGCAACC----- 566
Db 279 ThrProLysGluProAlaProThrAlaProLysLysProAlaProThrThrProLysGlu 298
QY 567 ---GGCCACCATCTCGCGGACGCAATATCGCGGAGCAATCTTCAAAGTCGGTGAGG 623
Db 299 ProAlaProThrThrProLysGluProAlaProThrThrLysGluProSerProThr 318
QY 624 ACACGGTTT-----TCGCGGACCTGGACATCTCGCCACCGCAGCGTCGTCGGCGAAG 677
Db 319 ThrProLysGluProAlaProThrThrLysSerAlaProThrThrLysGluPro 338
QY 678 TATCGACGACTTCGACCGCTACTGGGCAAGCCATTCGCCCAACAGCCGCGCATCA 737
Db 339 AlaProThrThrLysSerAla-----ProThrProLysGluPro 353
QY 738 TCCGAGCGGCAACATCGCAAGGTCCTCAAGCACTCGGATACACGACGAAACATCCA 797
Db 354 SerProThrThrLys-----GluProAlaProThrThrProLysGluPro 369
QY 798 -----GACACGGCTCTCGCTACCGCAACCGTGAACAGTCGCCCTCTACC 848
Db 370 AlaProThrThrProLysLysProAlaProThrThrProLysGluProAlaProThrThr 389
QY 849 AAAAAATACAGGAGCGATCGACTGGCAGAGGTCCAAACCCGCGCTGATCAGCG--- 905
Db 390 ProLysGluProAlaProThrThr-----LysLysProAlaProThrAlaPro 406
QY 906 -----ACACCCCTGCAAAAGGACTCGACCGCAGCGCCGCAACCGC 947
Db 407 LysGluProAlaProThrPro-----LysGluThrAlaProThrThrProLysLys 424
QY 948 CGATTGCCGGAGGTGCAAGACGCGCTCAACAGCCCGCAAAAGCGTCTACTCGTTT 1007
Db 425 LeuThrPro-----ThrThrProLysLysLeuAla----- 434
QY 1008 CACCCTATTTCCTCCTCAAAATCGGCAAGAGCTGCTACAGCGCAGCTGCAAACTGGTGCAGGAGC 1067
Db 435 -----ProThrThrProGluLysProAlaProThrThrProGlu-----GluLeu 449
QY 1068 GCATAGACGTTCGCTGACCACTCGCTACAGCGCAGCTGCGC----- 1118
Db 450 AlaProThrThrProGluLysProThr-----ProThrThrProGluLysPro 465
QY 1119 ---CCGTCAATTCGGCTAGCTCAATACCGAAACCGCTGCTCAAAGCCGCGCATCAAC 1175
Db 466 AlaProThrThrProLysAlaAla---AlaProAsnThrProLysGluProAla----- 482
QY 1176 TCTACGAGCTGCACCCACCATCGGTCGCCGCCACAAAGACAAAGCCCTGACGGCA 1235
Db 483 -----ProThrThrProLysGluProAlaProThrThr-----ProLys 495
QY 1236 GCTCCGTAACCGCTGCATGCCAAACCTTCATTGTGGACGCAACCGCATCTTCATCG 1295
Db 496 GluProAlaProThr---ThrProLysGluThrAlaProThrThrProLysGlyThrAla 514
QY 1296 GTCATTCAACCTCGACCCCGTTCGCGACGGCTCAATACCGAAATGGCGTGCATCG 1355
Db 515 ProThrThrLeuLysGluProAlaPro-----ThrThrProLys-----Lys 528
QY 1356 AAGCCCAATTCGAGACAGATGGAGCGCACCTCGCCGATACCACCGCAATACG 1415
Db 529 ProAlaProLysGlu-----LeuAlaProThrThrThrLysGluProThrSer 544
QY 1416 CTACCGCTTACCCCTCGCAACACACCGCGCTGCAATGGCAGCATCCCG---CCACCC 1472
Db 545 ThrThrSerAspLysProAlaProThrThrPro---LysGlyThrAlaProThrThrPro 563

QY 1473 GAAAAACCTACCGAAGCAACCGGAAG 1499
Db 564 LysGluProAlaProThrThrProLys 572
RESULT 9
US-11-169-232-84
; Sequence 84, Application US/11169232
; Publication No. US20060025570A1
; GENERAL INFORMATION:
; APPLICANT: Turner, Katherine
; Clark, Stephen C.
; Jacobs, Kenneth
; Hewick, Rodney M.
; Gesner, Thomas G.
; TITLE OF INVENTION: Megakaryocyte Stimulating Factors
; NUMBER OF SEQUENCES: 143
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Genetics Institute, Inc.
; STREET: 87 CambridgePark Drive
; CITY: Cambridge
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02140
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION NUMBER: US/11/169,232
; FILING DATE: 28-Jun-2005
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/10/124,557
; FILING DATE: 16-Apr-2002
; APPLICATION NUMBER: US 07/643,502
; FILING DATE: 18-JAN-1991
; APPLICATION NUMBER: US 07/546,114
; FILING DATE: 29-JUN-1990
; APPLICATION NUMBER: US 07/457,196
; FILING DATE: 29-DEC-1989
; APPLICATION NUMBER: US 07/390,901
; FILING DATE: 08-AUG-1989
; ATTORNEY/AGENT INFORMATION:
; NAME: Cseri, Luann
; REGISTRATION NUMBER: 31,822
; REFERENCE/DOCKET NUMBER: GI 5190
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)876-1170
; TELEFAX: (617)876-5851
; INFORMATION FOR SEQ ID NO: 84:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1022 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 84:
US-11-169-232-84
Alignment Scores: 5.48e-09 Length: 1022
Pred. No.: 247.50 Matches: 135
Score: 35.2% Conservative: 44
Percent Similarity: 26.5% Mismatches: 205
Query Match: 8.7% Indels: 125
DB: 11 Gaps: 29
US-10-665-990A-13 (1-1561) x US-11-169-232-84 (1-1022)
QY 144 CCAAACTGTCTCTGGCAACATCTGCAAAATCCGCAACCCCTCATACAAACGGGC 203
Db 213 ProThrProLysAlaGluThrThr-----LysGlyProAlaLeuThrThrProLys 230

QY 204 TATCGACATCTACC-----TGCTCGACGACCCCGACGAAGCCCTTGCGG 248
DB 231 GluProThrProThrProLysGluProAlaSerThrThrProLysGluProThrPro 250
QY 249 -----CCCGCGCGCCCTTATCGAATCTGCGCAACACAGCCCTCGATT 290
DB 251 ThrThrIleLysSerAlaProThrThrProLysGluProAlaProThrThrThrLysSer 270
QY 251 TGCAATACTACA-----TTTGGCGCAACGACATTT 320
DB 271 AlaProThrThrProLysGluProAlaProThrThrThrLysGluProAlaProThrThr 290
QY 321 CGGCGAGGCTGCTGTCAACTCTATGCTTCCGCGAGAACGCGCGCTGCGCGTACGCC 380
DB 291 ProLysGluProAlaProThrThrThrLysGluProAla----- 303
QY 381 TGCTTTGGACGACGAACACGCGCGGTGGACGATCTCTGCTCCCTTCGACAGCC 440
DB 304 -----ProThrThrThrLysSerAlaProThrThrProLysGluProAlaProThr 320
QY 441 ATCCCAATATCGAAGTGGCGCTGTCAACCCCTTCTGCTACGCAAAATGGCGGCACTCG 500
DB 321 ThrPro-----LysLysProAlaProThrThrProLysGluProAlaProThrProLys 339
QY 501 GCTACTCTGACGACTTCCCGCGCTCAACCGCGCATGCACAAACAAATCCTTTACCG--- 557
DB 340 GluProThrProThrProLysGluProAlaProThrThrThrLysGluProAlaProThr 359
QY 558 -----CGACAAAC----- 566
DB 360 ThrProLysGluProAlaProThrAlaProLysLysProAlaProThrThrProLysGlu 379
QY 567 ---GCGCCACCATCTCGCGGAGCGCAATATCGCGACGATACTTCAAAGTCGGTAGG 623
DB 380 ProAlaProThrProLysGluProAlaProThrThrThrLysGluProSerProThr 399
QY 624 ACACCGTTT-----TCGCCGACTGGACATCTCCGCCACCGCGAGCGCTCGCGGAAG 677
DB 400 ThrProLysGluProAlaProThrThrThrLysSerAlaProThrThrThrLysGluPro 419
QY 678 TATCGCAGACTTGCACCGTACTGGGCAAGCCATTCGCCCAACAGCCACGCGCATCA 737
DB 420 AlaProThrThrThrLysSerAla-----ProThrThrProLysGluPro 434
QY 738 TCCGAGCGGCAACATCGGCAAGGCTCTCAAGCACTCGGATACAGCAACGCAACATCCA 797
DB 435 SerProThrThrThrLys-----GluProAlaProThrThrProLysGluPro 450
QY 798 -----GACACGCGCTCTGCGCTACCGCGAAACCGTCGAACAGTCGCGCCCTCTACC 848
DB 451 AlaProThrThrProLysLysProAlaProThrThrProLysGluProAlaProThrThr 470
QY 849 AAAAAATACAGCGGAGCATCGACTGGCAGAGCGTCCAAACCGCGCTGATCAGCG--- 905
DB 471 ProLysGluProAlaProThrThr-----LysLysProAlaProThrAlaPro 487
QY 906 -----ACACCCCTGCAAAGGACTCGACCGCGCGCGCGCGCAACCGC 947
DB 488 LysGluProAlaProThrPro-----LysGluThrAlaProThrThrProLysLys 505
QY 948 CGATTGCGGAGGCTCGAAGCGCTCAACAGCCCGAAAAAGCGCTATCTCGGTTT 1007
DB 506 LeuThrPro-----ThrThrProGluLysLeuAla----- 515
QY 1008 CACCTATTTCCTCCCTACAAAATCCGGCAGACGAGCTGGCAAAACTGGTGAGGAGC 1067
DB 516 -----ProThrThrProGluLysProAlaProThrThrProGlu-----GluLeu 530
QY 1068 GCATAGAGGTACCGTCTCGCAACTCGCTACAGGCGACGAGGTGCGC----- 1118
DB 531 AlaProThrThrProGluGluProThr-----ProThrThrProGluGluPro 546

QY 1119 ---CGCTCATTCGCTACGTCAAATACCGAAACCGCTGCTCAAACCGCGCATCAAC 1175
DB 547 AlaProThrThrProLysAlaAla---AlaProAsnThrProLysGluProAla----- 563
QY 1176 TCTAGAGCTCAACCAACCAACCGCTCCCGCGCACAAAAGACAAAGGCGCTGACCGGCA 1235
DB 564 -----ProThrThrProLysGluProAlaProThrThr-----ProLys 576
QY 1236 GCTCGTAACAGCGCTGCATCGCAAAACCTTCATTGTGGAGCGGCAACGCAATCTTCATCG 1295
DB 577 GluProAlaProThr-----ThrProLysGluThrAlaProThrThrProLysGlyThrAla 595
QY 1296 GCTCATCAACCTCGACCCCGCTTCCGACGCGCTCAATACCGAAATGGCGCTCGTCATCG 1355
DB 596 ProThrThrLeuLysGluProAlaPro-----ThrThrProLys-----Lys 609
QY 1356 AAAGCCCAAAATCGCAGACAGATGGAGCGCACCTCGCGATACACACCGCAATACG 1415
DB 610 ProAlaProLysGlu-----LeuAlaProThrThrThrLysGluProThrSer 625
QY 1416 CTACCGCGTTCCTCGACAAACACACCGCGCTGCAATGGCAGATCCCG---CCAGCC 1472
DB 626 ThrThrSerAspLysProAlaProThrPro---LysGlyThrAlaProThrThrPro 644
QY 1473 GAAAAACCTACCCGACGAAACCCGAAG 1499
DB 645 LysGluProAlaProThrThrProLys 653

RESULT 10
US-11-169-232-74
; Sequence 74, Application US/11169232
; Publication No. US20060025570A1
; GENERAL INFORMATION:
; APPLICANT: Turner, Katherine
; Clark, Stephen C.
; Jacobs, Kenneth
; Hewick, Rodney M.
; Gesner, Thomas G.
; TITLE OF INVENTION: Megakaryocyte Stimulating Factors
; NUMBER OF SEQUENCES: 143
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genetics Institute, Inc.
; STREET: 87 Cambridgepark Drive
; CITY: Cambridge
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02140
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/11/169,232
; FILING DATE: 28-Jun-2005
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/10/124,557
; FILING DATE: 16-Apr-2002
; APPLICATION NUMBER: US 07/643,502
; FILING DATE: 18-JAN-1991
; APPLICATION NUMBER: US 07/546,114
; FILING DATE: 29-JUN-1990
; APPLICATION NUMBER: US 07/457,196
; FILING DATE: 29-DEC-1989
; APPLICATION NUMBER: US 07/390,901
; FILING DATE: 08-AUG-1989
; ATTORNEY/AGENT INFORMATION:
; NAME: Cseri, Luann
; REGISTRATION NUMBER: 31,822
; REFERENCE/DOCKET NUMBER: GI 5190
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 876-1170

Db	435	AlaProThrThrProLysLysProAlaProThrThrProLysGluProAlaProThrThr	454
QY	849	AAAAAATACAGACGGGACGCATCGACTGGCGACAGAGCTCCAAACCCGCGCTGATCAGCG	905
Db	455	ProLysGluProAlaProThrThrThr	471
QY	906	-----ACACCCCTGC AAAAGGACTCGACCGCGACCGCGCAAAACCGC	947
Db	472	LysGluProAlaProThrThrPro-----LysGluThrAlaProThrThrProLysLys	489
QY	948	CGATTCCCGGAGGCTGCAAGACGCGCTCAACACAGCCCGAAAAACGCTCTATCTGGTTT	1007
Db	490	LeuThrPro-----ThrThrProLysLysLeuAla	499
QY	1008	CACCCTATTTCGTCCTCAAAAATCGGGACAGACGCATCTGGCAAAACTGGTGCAGGACG	1067
Db	500	-----ProThrThrProLysLysProAlaProThrThrProGlu-----GluLeu	514
QY	1068	GCATAGACGTTACCGTCCTGCACCACTCGCTACAGCGGACCGACGTTGCCG	1118
Db	515	AlaProThrThrProLysGluProThr-----ProThrThrProLysGluPro	530
QY	1119	---CCGTCCATTCCGCTACGTCAAATACACGAAACCGCTGCTCAAAGCGGCATCAAAAC	1175
Db	531	AlaProThrThrProLysAlaAla---AlaProAsnThrProLysGluProAla	547
QY	1176	TCTACGAGCTGAACCAACCATCGCTCCGCGCACAAAAGACAAAGCGCTGACCGGCA	1235
Db	548	-----ProThrThrProLysGluProAlaProThrThr-----ProLys	560
QY	1236	GCTCCGTAACCGCTGCTGCAAAACCTTCATTGTGACGCGCAACGCATCTTCATCG	1295
Db	561	GluProAlaProThr---ThrProLysGluThrAlaProThrThrProLysGlyThrAla	579
QY	1296	GCTCATTCAACCTCGACCCCGTTCGCGAGGCTCAATACCGAAATGGGCGTCTCATCG	1355
Db	580	ProThrThrLeuLysGluProAlaPro-----ThrThrProLys-----Lys	593
QY	1356	AAAGCCCCAAAATCGCAGAACAGATGGAGCGCACCTCTCCGCGATACACACCGCAATACG	1415
Db	594	ProAlaProLysGlu-----LeuAlaProThrThrThrLysGluProThrSer	609
QY	1416	CCTACGCGCTTACCTCGCAAAACACACACCGCTGCAATGGCAGATCCCG---CCACCC	1472
Db	610	ThrThrSerAspLysProAlaProThrThrPro---LysGlyThrAlaProThrThrPro	628
QY	1473	GAANAACCTACCGACGACCGGANG	1499
Db	629	LysGluProAlaProThrThrProLys	637
RESULT 11			
US-11-169-232-58			
; Sequence 58, Application US/11169232			
; Publication No. US20060025570A1			
; GENERAL INFORMATION:			
; APPLICANT: Turner, Katherine			
; Clark, Stephen C.			
; Jacobs, Kenneth			
; Hewick, Rodney M.			
; Geaner, Thomas G.			
; TITLE OF INVENTION: Megakaryocyte Stimulating Factors			
; NUMBER OF SEQUENCES: 143			
; CORRESPONDENCE ADDRESS:			
; ADDRESSEE: Genetics Institute, Inc.			
; STREET: 87 CambridgePark Drive			
; CITY: Cambridge			
; STATE: Massachusetts			
; COUNTRY: U.S.A.			
; ZIP: 02140			
; COMPUTER READABLE FORM:			
; MEDIUM TYPE: Floppy disk			
; COMPUTER: IBM PC compatible			


```
QY 1176 TCTACGAGTGCACCCCAACCATGCGTCCCGCCACAAAGACAAAGAGCCCTGACCGGCA 1235
Db ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
682 -----ProThrThrProLysGluProAlaProThr-----ProLys 694
QY 1236 GTCGTAACGAGCTGCATGCAAAACCTTCATTTGTGGAGCGGCAACGATCTTCATCG 1295
Db ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
695 GluProAlaProThr---ThrProLysGluThrAlaProThrProLysGlyThrAla 713
QY 1296 GCTCATTCACCTGACCCCGCTTCCGACCGCTCAATACCGAAATGGCGTCGTCATCG 1355
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
714 ProThrThrLeuLysGluProAlaPro-----ThrThrProLys-----Lys 727
QY 1356 AAAGCCCCAAATCCAGACAGATGGAGCGACCTCGCGATACCAACCCGAATACG 1415
Db ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
728 ProAlaProLysGlu-----LeuAlaProThrThrThrLysGluProThrSer 743
QY 1416 CCTACCGGTACCTCGACAAACACACACCGCTGCAATGCGACGATCCCG---CCACCC 1472
Db ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
744 ThrThrSerAspLysProAlaProThrThrPro---LysGlyThrAlaProThrPro 762
QY 1473 GAAAAACCTACCGCAACCAACCGGAAG 1499
Db :|||: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
763 LysGluProAlaProThrThrProLys 771

RESULT 13
US-11-169-232-44
; Sequence 44, Application US/11169232
; Publication No. US20060025570A1
; GENERAL INFORMATION:
; APPLICANT: Turner, Katherine
; Clark, Stephen C.
; Jacobs, Kenneth
; Hewick, Rodney M.
; Gesner, Thomas G.
; TITLE OF INVENTION: Megakaryocyte Stimulating Factors
; NUMBER OF SEQUENCES: 143
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genetics Institute, Inc.
; STREET: 87 Cambridgepark Drive
; CITY: Cambridge
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02140
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/11/169,232
; FILING DATE: 28-Jun-2005
; CLASSIFICATION: <unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/10/124,557
; FILING DATE: 16-Apr-2002
; APPLICATION NUMBER: US 07/643,502
; FILING DATE: 18-JAN-1991
; APPLICATION NUMBER: US 07/546,114
; FILING DATE: 29-JUN-1990
; APPLICATION NUMBER: US 07/457,196
; FILING DATE: 29-DEC-1989
; APPLICATION NUMBER: US 07/390,901
; FILING DATE: 08-AUG-1989
; ATTORNEY/AGENT INFORMATION:
; NAME: Cseri, Luann
; REGISTRATION NUMBER: 31,822
; REFERENCE/DOCKET NUMBER: GI 5190
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 876-1170
; TELEFAX: (617) 876-5851
; INFORMATION FOR SEQ ID NO: 44:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1270 amino acids
```

```
;
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 44:
US-11-169-232-44
Alignment Scores:
Pred. No.: 5.65e-09 Length: 1270
Score: 247.50 Matches: 135
Percent Similarity: 35.2% Conservative: 44
Best Local Similarity: 26.5% Mismatches: 205
Query Match: 8.7% Indels: 125
DB: 11 Gaps: 29
US-10-665-990A-13 (1-1561) x US-11-169-232-44 (1-1270)
QY 144 CMAAACCTGCTCTCTCGAGCAACATCTGCAATTCGGCACACCCCTCATTAACACGGGC 203
Db ||| :|||: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
197 ProThrProLysAlaGluThrThr-----LysGlyProAlaLeuThrThrProLys 214
QY 204 TATCCGACATCTACC-----TGCTCGAGCAGCCCGACGAGCCCTTGGCG 248
Db ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
215 GluProThrProThrThrProLysGluProAlaSerThrThrProLysGluProThrPro 234
QY 249 -----CCGCGCGCGCTTATCGAATCTGCCGACACACAGCCCTCGATT 290
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
235 ThrThrIleLysSerAlaProThrThrProLysGluProAlaProThrThrThrLysSer 254
QY 291 TGCATTAATACTACA-----TTTGGCGCAACGACATTT 320
Db ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
255 AlaProThrThrProLysGluProAlaProThrThrThrLysGluProAlaProThrThr 274
QY 321 CCGCAGCGCTGCTTCAACCTCATGTACCTTGGCGAGAACGCGCGTGGCGGTACGCC 380
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
275 ProLysGluProAlaProThrThrThrLysGluProAla----- 287
QY 381 TGCTGTTGGACGACACACACGCGGTTGGACGATCTCTCTCGCCCTCGACAGCC 440
Db ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
288 -----ProThrThrThrLysSerAlaProThrThrProLysGluProAlaProThr 304
QY 441 ATCCCAATATCGAAGTGGCGCTGTTCAACCCCTTCTGCTACGAAATGGCGGCACTCG 500
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
305 ThrPro---LysLysProAlaProThrThrProLysGluProAlaProThrThrProLys 323
QY 501 GCTACCTGACCGACTTCCCGCGCTCAACCGCGCATGCGCAACAAATCTTTACCG--- 557
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
324 GluProThrProThrThrProLysGluProAlaProThrThrLysGluProAlaProThr 343
QY 558 -----CCGACAAAC----- 566
Db ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
344 ThrProLysGluProAlaProThrAlaProLysLysProAlaProThrThrProLysGlu 363
QY 567 ---GCGCCACCATCTCGCGGAGCGCAATATCGCGAGCAATCTTCAAAGTCGGTAGG 623
Db ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
364 ProAlaProThrThrProLysGluProAlaProThrThrThrLysGluProSerProThr 383
QY 624 ACACCGTTT-----TCGCGACCTGACATCTCGCCACCGGCGTGGCGCGCAAG 677
Db ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
384 ThrProLysGluProAlaProThrThrThrLysSerAlaProThrThrThrLysGluPro 403
QY 678 TATCGCACGACTTCGACCGCTACTGGGCAAGCCATTCGCGCCACACCGCCGCGATCA 737
Db ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
404 AlaProThrThrThrLysSerAla-----ProThrThrProLysGluPro 418
QY 738 TCGCGAGCGCAACATCGGCAAGGTTCTTCAAGACTCTCGGATCAACAGCAATATCCA 797
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
419 SerProThrThrThrLys-----GluProAlaProThrThrProLysGluPro 434
QY 798 -----GACACGCGCTCTCGCTACCGCGCAACCGTCGAACAGTCCCGCTCTACC 848
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
435 AlaProThrThrProLysLysProAlaProThrThrProLysGluProAlaProThrThr 454
QY 849 AAAAAATACAGACGGGACGCGATCGACTGGCAGAGCGTCCAAACCGCGCTGATCAGCG--- 905
```



```
Db 385 ThrProLysGluProAlaProThrAlaProLysLysProAlaProThrThrProLysGlu 404
QY 567 ---GGCCACCATACTCGCGGACGCAATATCGGACGCAATCTTCAAAGTCGGTGAGG 623
Db 405 ProAlaProThrThrProLysGluProAlaProThrThrThrLysGluProSerProThr 424
QY 624 ACACCGTTT-----TCGCCGACCTCGGACATCCTCGCCACCGCGGACGCGTCGCGCGAAG 677
Db 425 ThrProLysGluProAlaProThrThrThrLysSerAlaProThrThrThrLysGluPro 444
QY 678 TATCGACGACTTCACCGCTACTCGGCAAGCATTCCGCCACAAACCGCGACGCGCATCA 737
Db 445 AlaProThrThrThrLysSerAla-----ProThrThrProLysGluPro 459
QY 738 TCCGCGCGGCAACATCGGCAAGGTCTTCAAGCACTCGGATACAAACGACGCAAAATCA 797
Db 460 SerProThrThrThrLys-----GluProAlaProThrThrProLysGluPro 475
QY 798 -----GACACGCGCTCTCGCTACCGCGAAACGTCGAACAGTCGCCCTCTACC 848
Db 476 AlaProThrThrProLysProAlaProThrThrProLysGluProAlaProThrThr 495
QY 849 AAAAAATACACGGGACGCTCGACTCGCAGAGGTCCAAACCGCGCTGATCAGCG--- 905
Db 496 ProLysGluProAlaProThrThr-----LysLysProAlaProThrAlaPro 512
QY 906 -----ACACCCCTGCAAAAGGACTCGACCGCGCGACCGCGCGCAACCGC 947
Db 513 LysGluProAlaProThrPro-----LysGluThrAlaProThrThrProLysLys 530
QY 948 CGATTGCGGGAGGCTCAAGACGCGCTCAACACGCGGAAAGCGCTCTATCTGGTTT 1007
Db 531 LeuThrPro-----ThrThrProGluLysLeuAla----- 540
QY 1008 CACCTATTTCGCTCAAAAATCCGCGCACAGACGCTCGCAAAACGTCGTGAGGACG 1067
Db 541 -----ProThrThrProGluLysProAlaProThrThrProGlu-----GluLeu 555
QY 1068 GCATAGAGTTTACGCTCGTACCACTCGCTACAGGCGACGAGTTCGCG----- 1118
Db 556 AlaProThrThrProGluLupProThr-----ProThrThrProGluLupPro 571
QY 1119 ---CGTCCATTTCGGGTACGTAATAATACCGAAACCGCTCTCAAAGCGCGCAAC 1175
Db 572 AlaProThrThrProLysAlaAla---AlaProAsnThrProLysGluProAla----- 588
QY 1176 TCTAGAGCTCAACCCCAACATCGCTCGCCCGCCGCAAAAAGACAAAGCGCTGACGCGCA 1235
Db 589 -----ProThrThrProLysGluProAlaProThrThr-----ProLys 601
QY 1236 GCTCGTAACGAGCTGCATGCCAAACCTTCATTGTGGACGCGCAACGCACTTCATCG 1295
Db 602 GluProAlaProThr-----ThrProLysGluThrAlaProThrThrProLysGlyThrAla 620
QY 1296 GCTCATTTCAACCTCGACCCCGTTTCGCGCACGCTCAATACCGAAATGGCGCTCATCG 1355
Db 621 ProThrThrLeuLysGluProAlaPro-----ThrThrProLys-----Lys 634
QY 1356 AAAGCCCAAAATCGCAACAGATGGAGCGCACCTCGCGCGATACCAACCCGAATACG 1415
Db 635 ProAlaProLysGlu-----LeuAlaProThrThrThrLysGluProThrSer 650
QY 1416 CCTACCGGTTACCTCGACAAACACACGCGCTGCAATGCGCAGATCCCG---CCACCC 1472
Db 651 ThrThrSerAspLysProAlaProThrThrPro-----LysGlyThrAlaProThrThrPro 669
QY 1473 GAAAAACCTACCCGAAGAACCCGAAG 1499
Db 670 LysGluProAlaProThrThrProLys 678
; Publication No. US20060025570A1
; GENERAL INFORMATION:
; APPLICANT: Turner, Katherine
; Clark, Stephen C.
; Jacobs, Kenneth
; Hewick, Rodney M.
; Gesner, Thomas G.
; TITLE OF INVENTION: Megakaryocyte Stimulating Factors
; NUMBER OF SEQUENCES: 143
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genetics Institute, Inc.
; STREET: 87 CambridgePark Drive
; CITY: Cambridge
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02140
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/11/169,232
; FILING DATE: 28-Jun-2005
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/10/124,557
; FILING DATE: 16-Apr-2002
; APPLICATION NUMBER: US 07/643,502
; FILING DATE: 18-JAN-1991
; APPLICATION NUMBER: US 07/546,114
; FILING DATE: 29-JUN-1990
; APPLICATION NUMBER: US 07/457,196
; FILING DATE: 29-DEC-1989
; APPLICATION NUMBER: US 07/390,901
; FILING DATE: 08-AUG-1989
; ATTORNEY/AGENT INFORMATION:
; NAME: Cserr, Luann
; REGISTRATION NUMBER: 31,822
; REFERENCE/DOCKET NUMBER: GI 5190
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 876-1170
; TELEFAX: (617) 876-5851
; INFORMATION FOR SEQ ID NO: 142:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1313 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 142:
US-11-169-232-142
Alignment Scores:
Pred. No.: 5,67e-09 Length: 1313
Score: 247.50 Matches: 135
Percent Similarity: 35.2% Conservative: 44
Best Local Similarity: 26.5% Mismatches: 205
Query Match: 8.7% Indels: 125
DB: 11 Gaps: 29
US-10-665-990A-13 (1-1561) x US-11-169-232-142 (1-1313)
QY 144 CCAAACTGCTCTCGGACAAATCTGCAATCCGACACCCCTCATTAACACGGGC 203
Db 240 ProThrProLysAlaGluThrThrThr-----LysGlyProAlaLeuThrThrProLys 257
QY 204 TATCCGACATCTACC-----TGCTCGACGACCCCGACGAAGCCCTTCGCG 248
Db 258 GluProThrProThrThrProLysGluProAlaSerThrThrProLysGluProThrPro 277
QY 249 -----CCCGCGCGCCCTTATCGAATCTGCGAACAACAGCCGCTCGATT 290
Db 278 ThrThrIleLysSerAlaProThrThrProLysGluProAlaProThrThrThrLysSer 297
; Sequence 142, Application US/11169232
```

RESULT 15

US-11-169-232-142

; Sequence 142, Application US/11169232

QY 291 TGCAATACTACA-----TTTGGCGCAACGACATT 320
Db 298 AlaProThrThrProLysGluProAlaProThrThrLysGluProAlaProThrThr 317
QY 321 CGGGCAGGCTGCTGTTTCAACCTCATGTACCTTGGCGCAGACGGCGTGGCGTACGCC 380
Db 318 ProLysGluProAlaProThrThrThrLysGluProAla----- 330
QY 381 TGCTGTTGGACGACAAACACGCGCGGTGGAGCATCTCTGCTCGCCCTCGACAGCC 440
Db 331 -----ProThrThrThrLysSerAlaProThrThrProLysGluProAlaProThr 347
QY 441 ATCCCAATATCGAAGTGGCGCTTCAACCCCTTCTGCTAGCGAAATGGCGGCGACTCG 500
Db 348 ThrPro-----LysLysProAlaProThrThrProLysGluProAlaProThrThrProLys 366
QY 501 GCTACCTGACGACTTCCCGCCTCAACCGCGCATGCACAAACAAATCTTTACCG--- 557
Db 367 GluProThrProThrThrProLysGluProAlaProThrThrLysGluProAlaProThr 386
QY 558 -----CCGACAAC----- 566
Db 387 ThrProLysGluProAlaProThrAlaProLysLysProAlaProThrThrProLysGlu 406
QY 567 ---GGCCACCATCTCGCGGAGCCATATCGCGGACGAAATCTTCAAAATCGGTGAGG 623
Db 407 ProAlaProThrThrProLysGluProAlaProThrThrThrLysGluProSerProThr 426
QY 624 ACACCGTTT-----TCGCGACCTGGACATCTCGCCACCGGACGCTGTCGCGGAAG 677
Db 427 ThrProLysGluProAlaProThrThrThrLysSerAlaProThrThrThrLysGluPro 446
QY 678 TATCCACGACTTCGACCGCTACTGGCAAGCATTCGCGCCACAAACGCCACGCGCATCA 737
Db 447 AlaProThrThrThrLysSerAla-----ProThrThrProLysGluPro 461
QY 738 TCCGAGCGGCAACATCGGCAAGGTCTTCAAGCACTCGGATACAAACGACGAAACATCCA 797
Db 462 SerProThrThrThrLys-----GluProAlaProThrThrProLysGluPro 477
QY 798 -----GACACGCGTCTCGCTACCGGAAACCGTGAACAGTCGCCCTCTACC 848
Db 478 AlaProThrThrProLysLysProAlaProThrThrProLysGluProAlaProThrThr 497
QY 849 AAAAAATACAGACGGGACGATCGACTGGCAGAGCGTCCAAACCGCCTGATCAGCG--- 905
Db 498 ProLysGluProAlaProThrThr-----LysLysProAlaProThrAlaPro 514
QY 906 -----ACACCCCTGCAAAAGGACTCGACCGCGACCGCGCAACCGC 947
Db 515 LysGluProAlaProThrThrPro-----LysGluThrAlaProThrThrProLysLys 532
QY 948 CGATTGCGGGAGGTGCAAGACGCGCTCAAAACAGCCGMAAAAGCGTCTATCTGGTTT 1007
Db 533 LeuThrPro-----ThrThrProGluLysLeuAla----- 542
QY 1008 CACCCTATTTCGCTCAAAATCCGGCAGACGCACTGGCAAAACCTGGTGCAGGACG 1067
Db 543 -----ProThrThrProGluLysProAlaProThrThrProGlu-----GluLeu 557
QY 1068 GCATAGACGTTACCGTCTGACCACTCGCTACAGCGGACCGAGCTTGGCG----- 1118
Db 558 AlaProThrThrProGluProThr-----ProThrThrProGluGluPro 573
QY 1119 ---CCGTCCATTCCGGCTACGTCAAATACCGAAACCGGTGCTCAAAACCGGCATCAAC 1175
Db 574 AlaProThrThrProLysAlaAla---AlaProAsnThrProLysGluProAla----- 590
QY 1176 TCTACGAGTGCACCCCAACCATCGCTCCCGCCACAAAGAGCCCTGACCGGCA 1235
Db 591 -----ProThrThrProLysGluProAlaProThrThr-----ProLys 603

QY 1236 GCTCGTAACAGCCTGCATGCGCAAAACCTTCTATTGGAGCGCAACGCAATCTTCATCG 1295
Db 604 GluProAlaProThr-----ThrProLysGluThrAlaProThrThrProLysGlyThrAla 622
QY 1296 GCTCATTAACACTCGACCCCGTTCCGACGCGCTCAATACCGAAATGGCGTGGTCAATCG 1355
Db 623 ProThrThrLeuLysGluProAlaPro-----ThrThrProLys-----Lys 636
QY 1356 AAAGCCCAAAATCGCAGAACAGATGGAGCGCACCTCGCGCGATACCAACCCGAAATACG 1415
Db 637 ProAlaProLysGlu-----LeuAlaProThrThrThrLysGluProThrSer 652
QY 1416 CCTACCGGTTACCTCGACAAACACACACGACGCGCTGCAATGGCAGATCCCG---CCACCC 1472
Db 653 ThrThrSerAspLysProAlaProThrThrPro-----LysGlyThrAlaProThrThrPro 671
QY 1473 GAAAAACCTACCCGAAACGAAACCCGAAG 1499
Db 672 LysGluProAlaProThrThrProLys 680

RESULT 16
US-11-169-232-50
; Sequence 50, Application US/11169232
; Publication No. US20060025570A1
; GENERAL INFORMATION:
; APPLICANT: Turner, Katherine
; Clark, Stephen C.
; Jacobs, Kenneth
; Hewick, Rodney M.
; Gesner, Thomas G.
; TITLE OF INVENTION: Megakaryocyte Stimulating Factors
; NUMBER OF SEQUENCES: 143
; CORRESPONDENCE ADDRESS:
; ADDRESS: Genetics Institute, Inc.
; STREET: 87 CambridgePark Drive
; CITY: Cambridge
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02140
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/11/169,232
; FILING DATE: 28-Jun-2005
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/10/124,557
; FILING DATE: 16-Apr-2002
; APPLICATION NUMBER: US 07/643,502
; FILING DATE: 18-JAN-1991
; APPLICATION NUMBER: US 07/546,114
; FILING DATE: 29-JUN-1990
; APPLICATION NUMBER: US 07/457,196
; FILING DATE: 29-DEC-1989
; APPLICATION NUMBER: US 07/390,901
; FILING DATE: 08-AUG-1989
; ATTORNEY/AGENT INFORMATION:
; NAME: Csert, Luann
; REGISTRATION NUMBER: 31,822
; REFERENCE/DOCKET NUMBER: GI 5190
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)876-1170
; TELEFAX: (617)876-5851
; INFORMATION FOR SEQ ID NO: 50:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1314 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 50:

US-11-169-232-50

Alignment Scores:

Pred. No.: 5,67e-09 Length: 1314
 Score: 247.50 Matches: 135
 Percent Similarity: 35.2% Conservative: 44
 Best Local Similarity: 26.5% Mismatches: 205
 Query Match: 8.7% Indels: 125
 DB: 11 Gaps: 29

US-10-665-990A-13 (1-1561) x US-11-169-232-50 (1-1314)

QY 144 CCAACCTGCTCTCTGACGAACATCTGCAAAATCCGACACACCCCTCATACAAACGGCG 203
 DB 241 ProThrProLysAlaGluThrThr-----LysGlyProAlaLeuThrProLys 258
 QY 204 TATCGACATCTACC-----TGCTCGACGACCCCGACGAAAGCCCTTGCGC 248
 DB 259 GluProThrProThrProLysGluProAlaSerThrThrProLysGluProThrPro 278
 QY 249 -----CCCGCGCGCCCTTATCGAATCTGCGAATCGCACACAGCCCTCGATT 290
 DB 279 ThrThrLysSerAlaProThrThrProLysGluProAlaProThrThrLysSer 298
 QY 291 TGCATATCTACA-----TTTGGCGCAACGACATTT 320
 DB 299 AlaProThrThrProLysGluProAlaProThrThrLysGluProAlaProThrThr 318
 QY 321 CGGCGAGGCTGCTTCAACCTCATGTACCTTGCGGAGAGCGGGCGGTGCGGTACGCC 380
 DB 319 ProLysGluProAlaProThrThrLysGluProAla----- 331
 QY 381 TGCTGTTGACGACAAACACGCGGGTTGGAGCATCTCTGCTGCGCCCTCGACAGCC 440
 DB 332 -----ProThrThrLysSerAlaProThrThrProLysGluProAlaProThr 348
 QY 441 ATCCCAATATCGAAGTGGCTGTTCACCCCTTCTGCTACGCAAAATGGCGCGCACTCG 500
 DB 349 ThrPro---LysLysProAlaProThrThrProLysGluProAlaProThrThrProLys 367
 QY 501 GCTACCTGACGACTTCCCGCGCTCAACCGCGGATGCAACAAATCCTTTACCG--- 557
 DB 368 GluProThrProThrProLysGluProAlaProThrThrLysGluProAlaProThr 387
 QY 558 -----CCGCAACAC----- 566
 DB 388 ThrProLysGluProAlaProThrAlaProLysLysProAlaProThrThrProLysGlu 407
 QY 567 ---CGGCCACCATCTCGGCGGACGCAATATCGGCGACGATACTTCAAAAGTCGGTGAGG 623
 DB 408 ProAlaProThrThrProLysGluProAlaProThrThrLysGluProSerProThr 427
 QY 624 ACACGGTT-----TCGCCACCTGGACATCTCGCCACCGGAGGTCGTGCGGCGAAG 677
 DB 428 ThrProLysGluProAlaProThrThrLysSerAlaProThrThrLysGluPro 447
 QY 678 TATCGCAGACTTCGACCGCTACTGGGCAAGCCATTCGCGCCCAACGACGCGCATCA 737
 DB 448 AlaProThrThrLysSerAla-----ProThrThrProLysGluPro 462
 QY 738 TCCGACGGCAACATCGGCAAGGCTTCAAGCACTCGGATACAAACGACGAAACATCCA 797
 DB 463 SerProThrThrLys-----GluProAlaProThrThrProLysGluPro 478
 QY 798 -----GACACGGCTCTCGCTACCGCGAAACGGTCGAACAGTCGCGCTCTAC 848
 DB 479 AlaProThrThrProLysLysProAlaProThrThrProLysGluProAlaProThrThr 498
 QY 849 AAAAAATACACGCGGACGCTGACCTGGCAGAGCGTCCAAACCGCTGATACAGC--- 905
 DB 499 ProLysGluProAlaProThrThr-----LysLysProAlaProThrAlaPro 515
 QY 906 -----ACACCCCTGCAAAAGGACTCGACCGCGCGCGCGCAACCGC 947

Db 516 LysGluProAlaProThrPro-----LysGluThrAlaProThrProLysLys 533
 QY 948 CGATTGCGGGAGGCTGCAAGACGCGCTCAACACGCGCAAAAGCGCTCTATCTGGTTT 1007
 Db 534 LeuThrPro-----ThrThrProGluLysLeuAla----- 543
 QY 1008 CACCTATTTCGCTCCCTACAAAATCCGCGACAGACGACTGGCAAACTGCTGCGAGGACG 1067
 Db 544 -----ProThrThrProGluLysProAlaProThrThrProGlu-----GluLeu 558
 QY 1068 GCATAGAGCTTACGCTCTGACCAACTCGCTACAGGCGACGAGTGTGCGC----- 1118
 Db 559 AlaProThrThrProGluProThr-----ProThrThrProGluGluPro 574
 QY 1119 ---CCGTCCATTCCGGCTACGTCAAATACCGAAACCGCTCTCAAGCCGCGCTCAAAAC 1175
 Db 575 AlaProThrThrProLysAlaAla---AlaProsnThrProLysGluProAla----- 591
 QY 1176 TCTACGAGCTGCAACCCCAACCATCGCTCCCGCCACAAAAGACAAAGGCGCTGACCGCA 1235
 Db 592 -----ProThrThrProLysGluProAlaProThr-----ProLys 604
 QY 1236 GCTCCGTAACGAGCTGCTGCAAAACCTTTCATTGTGAGCGCAAGCGCATCTTCATCG 1295
 Db 605 GluProAlaProThr---ThrProLysGluThrAlaProThrThrProLysGlyThrAla 623
 QY 1296 GCTCATTCAACCTGACCCCGTTCGCGACGGCTCAATACGAAATGGGCGCTGCTCATCG 1355
 Db 624 ProThrThrLeuLysGluProAlaPro-----ThrThrProLys-----Lys 637
 QY 1356 AAAGCCCAAAATCGCAGAACAGATGAGCGGCGCCCTCGCGATACACACCCGAATACG 1415
 Db 638 ProAlaProLysGlu-----LeuAlaProThrThrThrLysGluProThrSer 653
 QY 1416 CCTACCGCTTACCTCGCAAAACACCGCTTCAATGSCAGATCCCG---CCACCC 1472
 Db 654 ThrThrSerAspLysProAlaProThrPro---LysGlyThrAlaProThrThrPro 672
 QY 1473 GAAAAACCTACCGAAGCAACCCGAAG 1499
 Db 673 LysGluProAlaProThrThrProLys 681

RESULT 17
 US-11-169-232-46
 ; Sequence 46, Application US/11169232
 ; Publication No. US20060025570A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Turner, Katherine
 ; Clark, Stephen C.
 ; Jacobs, Kenneth
 ; Hewick, Rodney M.
 ; Gesner, Thomas G.
 ; TITLE OF INVENTION: Megakaryocyte Stimulating Factors
 ; NUMBER OF SEQUENCES: 143
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Genetics Institute, Inc.
 ; STREET: 87 CambridgePark Drive
 ; CITY: Cambridge
 ; STATE: Massachusetts
 ; COUNTRY: U.S.A.
 ; ZIP: 02140
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/11/169,232
 ; FILING DATE: 28-Jun-2005
 ; CLASSIFICATION: <Unknown>
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US/10/124,557

```

; FILING DATE: 16-Apr-2002
; APPLICATION NUMBER: US 07/643,502
; FILING DATE: 18-JAN-1991
; APPLICATION NUMBER: US 07/546,114
; FILING DATE: 29-JUN-1990
; APPLICATION NUMBER: US 07/457,196
; FILING DATE: 29-DEC-1989
; APPLICATION NUMBER: US 07/390,901
; FILING DATE: 08-AUG-1989
; ATTORNEY/AGENT INFORMATION:
; NAME: Cseri, Luann
; REGISTRATION NUMBER: 31,822
; REFERENCE/DOCKET NUMBER: GI 5190
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)876-1170
; TELEFAX: (617)876-5851
; INFORMATION FOR SEQ ID NO: 46:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1320 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 46:
US-11-169-232-46

Alignment Scores:
Pred. No.: 5,67e-09 Length: 1320
Score: 247.50 Matches: 135
Percent Similarity: 35.2% Conservative: 44
Best Local Similarity: 26.5% Mismatches: 205
Query Match: 8.7% Indels: 125
DB: 11 Gaps: 29

US-10-665-990A-13 (1-1561) x US-11-169-232-46 (1-1320)
Qy 144 CCAAACTCTCTCTCGGACAACTCTGCAAAATCCGGCACACCCCTCATAAACAACGGGC 203
Db ||| :: |||||::: ||| ||| ||| |||
Qy 247 ProThrProLysAlaGluThrThr-----LysGlyProAlaLeuThrThrProLys 264
Db ||| ||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 204 TATCCGACATCTACG-----TGCTCGAGACCCCGCACGAAGCCCTTGCCG 248
Db ||||| ||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 265 GluProThrProThrThrProLysGluProAlaSerThrThrProLysGluProThrPro 284
Db ||||| ||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 249 -----CCGCGCGCGCCCTTATCGAATCTCGGAACACAGCGCTCGATT 290
Db ||||| ||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 285 ThrThrIleLysSerAlaProThrThrProLysGluProAlaProThrThrThrLysSer 304
Db ||||| ||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 291 TGCATACTACA-----TTTGGCGGAACGACATTT 320
Db ||||| ||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 305 AlaProThrProLysGluProAlaProThrThrThrLysGluProAlaProThrThr 324
Db ||||| ||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 321 CCGCGACGGCTGTGTTCAACCTCATGTACCTTTCGCGACGAACGGCGTGGCGGTACGCC 380
Db ||||| ||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 325 ProLysGluProAlaProThrThrThrLysGluProAla----- 337
Db ||||| ||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 381 TGCTGTTGGACGACAAACAGCGCGCGGTTGGACGATCTCTGCTCGCCTCGACAGCC 440
Db ||||| ||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 338 -----ProThrThrThrLysSerAlaProThrThrProLysGluProAlaProThr 354
Db ||||| ||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 441 ATCCCAATATCAAGTGGCGCTGTTCAACCCCTTCGCTACGCAAAATGGCGGCACTCG 500
Db ||||| ||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 355 ThrPro---LysLysProAlaProThrThrProLysGluProAlaProThrProLys 373
Db ||||| ||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 501 GCTACTGACCGACTTCCCGCCTCAACCGCGCATGACAAACAAATCTTTACCG- 557
Db ||||| ||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 374 GluProThrProThrThrProLysGluProAlaProThrThrLysGluProAlaProThr 393
Db ||||| ||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 558 -----CCGACAACC----- 566
Db ||||| ||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 394 ThrProLysGluProAlaProThrAlaProLysLysProAlaProThrThrProLysGlu 413
Db ||||| ||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 567 ---GCGCCACCACTACGCGGACGACAAATCGCGGACGAATACTTCAAAGTCGGTGAGG 623
Db ||||| ||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```



```
Db 574 LeuThrPro-----:::||||:::| 583
Qy 1008 CACCCTATTTCGTCCCTACAAATCCGCACAGCGCACTGGCAAAACTGTGTGAGGACG 1067
Db 584 -----ProThrThrProGluLysProAlaProThrThrProGlu-----GluLeu 598
Qy 1068 GCATAGAGCTTACCGTCTCGCAACTCGCTACAGCGCAGCGAGCTTGCCG-----1118
Db 599 AlaProThrThrProGluLysProThr-----ProThrThrProGluLysPro 614
Qy 1119 ---CGTCCATTCCGGTACGTCAATAACCGAAACCGCTGCTCAAGCCGCGCATCAAC 1175
Db 615 AlaProThrThrProLysAlaAla---AlaProAsnThrProLysGluProAla-----631
Qy 1176 TCTAGAGCTCAACCCCAACCATCGCTCCCGCCCGCCACAAAGAGCAAGCGCTGACCGCA 1235
Db 632 -----ProThrThrProLysGluProAlaProThr-----ProLys 644
Qy 1236 GCTCGTAACGAGCTGCATGCCAAACCTTCATTGTGGAGCGCAACGCAATCTTCATCG 1295
Db 645 GluProAlaProThr---ThrProLysGluThrAlaProThrThrProLysGlyThrAla 663
Qy 1296 GCTCATTCACCTGACCCCGTTCGCGACGCTCAATACCGAAATGGCGTGTCTATCG 1355
Db 664 ProThrThrLeuLysGluProAlaPro-----ThrThrProLys-----Lys 677
Qy 1356 AAAGCCCAAAATCGCAGAACAGATGGAGCGCACCTCGCGATACCACACCCGGAATACG 1415
Db 678 ProAlaProLysGlu-----LeuAlaProThrThrThrLysGluProThrSer 693
Qy 1416 CCTACCGGTTTACCTCGACAAACACACACCGCTGCAATGGCAGATCCCG-----CCACCC 1472
Db 694 ThrThrSerAspLysProAlaProThrThrPro---LysGlyThrAlaProThrThrPro 712
Qy 1473 GAAAAACCTACCGAAGCAAGCCGGAAG 1499
Db 713 LysGluProAlaProThrThrProLys 721

RESULT 20
US-11-169-232-40
; Sequence 40, Application US/11169232
; Publication No. US20060025570A1
; GENERAL INFORMATION:
; APPLICANT: Turner, Katherine
; Jacobs, Kenneth
; Hewick, Rodney M.
; Gesner, Thomas G.
; TITLE OF INVENTION: Megakaryocyte Stimulating Factors
; NUMBER OF SEQUENCES: 143
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genetics Institute, Inc.
; STREET: 87 Cambridgepark Drive
; CITY: Cambridge
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02140
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/11/169,232
; FILING DATE: 28-Jun-2005
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/10/124,557
; FILING DATE: 16-Apr-2002
; APPLICATION NUMBER: US 07/643,502
; FILING DATE: 18-JAN-1991
; APPLICATION NUMBER: US 07/546,114
```

```
;
; FILING DATE: 29-JUN-1990
; APPLICATION NUMBER: US 07/457,196
; FILING DATE: 29-DEC-1989
; APPLICATION NUMBER: US 07/390,901
; FILING DATE: 08-AUG-1989
; ATTORNEY/AGENT INFORMATION:
; NAME: Cseri, Luann
; REGISTRATION NUMBER: 31,822
; REFERENCE/DOCKET NUMBER: GI 5190
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)876-1170
; TELEFAX: (617)876-5851
; INFORMATION FOR SEQ ID NO: 40:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1361 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 40:
US-11-169-232-40

Alignment Scores:
Pred. No.: 5,7e-09 Length: 1361
Score: 247.50 Matches: 135
Percent Similarity: 35.2% Conservative: 44
Best Local Similarity: 26.5% Mismatches: 205
Query Match: 8.7% Indels: 125
DB: 11 Gaps: 29

US-10-665-990A-13 (1-1561) x US-11-169-232-40 (1-1361)
Qy 144 CCAACCTGCTCTCTGACAGCAATCCTGCAAAATCGGCACACCCCTCATAAACACGGC 203
Db 288 ProThrProLysAlaGluThrThr-----LysGlyProAlaLeuThrThrProLys 305
Qy 204 TATCCGACATCTACC-----TGCTCGACGACCCCGCCAGAGCCCTTGCCG 248
Db 306 GluProThrProThrProLysGluProAlaSerThrThrProLysGluProThrPro 325
Qy 249 -----CCGCGCGCCCTTATCGAATCTGCGGAACACAGCCCTCGATT 290
Db 326 ThrThrIleLysSerAlaProThrThrProLysGluProAlaProThrThrLysSer 345
Qy 291 TGCAATACTACA-----TTTGGCGCAACGACATTT 320
Db 346 AlaProThrThrProLysGluProAlaProThrThrThrLysGluProAlaProThr 365
Qy 321 CCGGAGGCTGTGTTCAACCTCATGTACTTCCCGCAGAACGCGCGTGGCGGTACGCC 380
Db 366 ProLysGluProAlaProThrThrLysGluProAla-----378
Qy 381 TGCTGTGGACGACAAACACACGCGGGTGGACGATCTCTGCTCGCCCTCGACAGCC 440
Db 379 -----ProThrThrLysSerAlaProThrThrProLysGluProAlaProThr 395
Qy 441 ATCCCAATATCGAAGTGGCGCTGTCAACCCCTTGGTCTCGTCAAGAAATGCGCGCACTCG 500
Db 396 ThrPro---LysLysProAlaProThrThrProLysGluProAlaProThrThrProLys 414
Qy 501 GCTACTGACCGACTTCCCGCGCTCAACCGCGCATGACAAACAAATCCTTTACCG---557
Db 415 GluProThrProThrProLysGluProAlaProThrThrLysGluProAlaProThr 434
Qy 558 -----CCGCAACCC-----566
Db 435 ThrProLysGluProAlaProThrAlaProLysLysProAlaProThrThrProLysGlu 454
Qy 567 ---GCGCCACCATCTCGGCGGACGCAATATCGGCAGCAATATCTCAAAGTCGGTGAGG 623
Db 455 ProAlaProThrThrProLysGluProAlaProThrThrThrLysGluProSerProThr 474
Qy 624 ACACCGTTT-----TCGCGACCTGGACATCTCGCCACCGGACGCGTGTGGCGAAG 677
||||| :||||| ||| ||| :|||
```

```
Db 475 ThrProLysGluProAlaProThrThrThrLysSerAlaProThrThrThrLysGluPro 494
Qy 678 TATCGACGACTTCGACCGCTACTGGCAAGCCATTCCGCCCAACACCCGCGCATCA 737
Db 495 AlaProThrThrThrLysSerAla-----ProThrThrProLysGluPro 509
Qy 738 TCCGAGCGGCAACATCGCAAGGCTCTCAAGCACTGGATACAAACGACGAAACATCCA 797
Db 510 SerProThrThrThrLys-----GluProAlaProThrThrProLysGluPro 525
Qy 798 -----GACACGCGCTCTCGCTACCGGAAACGTCGAACAGTCGCGCCCTCTACC 848
Db 526 AlaProThrThrProLysProAlaProThrThrProLysGluProAlaProThrThr 545
Qy 849 AAAAAATACAGCGGACGATCGATGCGCAGAGCGTCCAAACCCGCTGTACGCG--- 905
Db 546 ProLysGluProAlaProThrThr-----LysLysProAlaProThrAlaPro 562
Qy 906 -----ACACCCCTGCNAAGGACTCGACCGCGACCGCCGCAACCGC 947
Db 563 LysGluProAlaProThrThrPro-----LysGluThrAlaProThrProLysLys 580
Qy 948 CGATTGCCGGGAGGCTGCAAGCGCGCTCAACAGCCGCAAAAGCGTCTATCTGTTT 1007
Db 581 LeuThrPro-----ThrThrProGluLysLeuAla----- 590
Qy 1008 CACCTATTTCGCTCAAAATCCGCGCAGACGCACTGCAAACTGGTGCAGGACG 1067
Db 591 -----ProThrThrProGluLysProAlaProThrThrProGlu-----GluLeu 605
Qy 1068 GCATAGAGTTACGCTCGACCACTCGCTACAGCGACCGACGCTTGCCG----- 1118
Db 606 AlaProThrThrProGluGluProThr-----ProThrThrProGluGluPro 621
Qy 1119 ---CGTCCATTCCGCTACGTCAAATACCGAAACCGCTGCTCAAGCCGCGCATCAAC 1175
Db 622 AlaProThrThrProLysAlaAla---AlaProAsnThrProLysGluProAla----- 638
Qy 1176 TCTACGAGTGCACCAACCAACCATGCGTCCCGCCGCAACAAAGACGCGTGCAGCA 1235
Db 639 -----ProThrThrProLysGluProAlaProThrThr-----ProLys 651
Qy 1236 GCTCGTAACGACGCTGCATGCAAACTTCATTGTGACGCGCAACCGCATCTTCATCG 1295
Db 652 GluProAlaProThr-----ThrProLysGluThrAlaProThrThrProLysGlyThrAla 670
Qy 1296 GCTATTCAACCTCGACCCCGTCCGACGCGTCAATACCGAAATGGCGTGCATCG 1355
Db 671 ProThrThrLeuLysGluProAlaPro-----ThrThrProLys-----Lys 684
Qy 1356 AAGGCCCAAAATCGCAGAACAGATGGAGCGCACCTCGCCGATACCACACCGAATACG 1415
Db 685 ProAlaProLysGlu-----LeuAlaProThrThrThrLysGluProThrSer 700
Qy 1416 CCTACCGCGTTACCTCGACAAACACACCGCTGCAATGGACGATCCG---CCACCC 1472
Db 701 ThrThrSerAspLysProAlaProThrThrPro---LysGlyThrAlaProThrThrPro 719
Qy 1473 GAAAAACCTTACCCGACGACCGCAAG 1499
Db 720 LysGluProAlaProThrThrProLys 728
```

RESULT 21

```
US-11-169-232-52
; Sequence 52, Application US/11169232
; Publication No. US20060025570A1
; GENERAL INFORMATION:
; APPLICANT: Turner, Katherine
; Clark, Stephen C.
; Jacobs, Kenneth
; Hewick, Rodney M.
; Gesner, Thomas G.
; TITLE OF INVENTION: Megakaryocyte Stimulating Factors
```

```
; NUMBER OF SEQUENCES: 143
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genetics Institute, Inc.
; STREET: 87 CambridgePark Drive
; CITY: Cambridge
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02140
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/11/169,232
; FILING DATE: 28-Jun-2005
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/10/124,557
; FILING DATE: 16-Apr-2002
; APPLICATION NUMBER: US 07/643,502
; FILING DATE: 18-JAN-1991
; APPLICATION NUMBER: US 07/546,114
; FILING DATE: 29-JUN-1990
; APPLICATION NUMBER: US 07/457,196
; FILING DATE: 29-DEC-1989
; APPLICATION NUMBER: US 07/390,901
; FILING DATE: 08-AUG-1989
; ATTORNEY/AGENT INFORMATION:
; NAME: Cserr, Luann
; REGISTRATION NUMBER: 31,822
; REFERENCE/DOCKET NUMBER: GI 5190
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)876-1170
; TELEFAX: (617)876-5851
; INFORMATION FOR SEQ ID NO: 52:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1363 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 52:
US-11-169-232-52
Alignment Scores:
Pred. No.: 5,7e-09 Length: 1363
Score: 247.50 Matches: 135
Percent Similarity: 35.2% Conservative: 44
Best Local Similarity: 26.5% Mismatches: 205
Query Match: 8.7% Indels: 125
DB: 11 Gaps: 29
US-10-665-990A-13 (1-1561) x US-11-169-232-52 (1-1363)
Qy 144 CCAACACCTGCTCTCTGGCAACATCTCTGCAATCCGCGCACACCCCTCATACAACGGGC 203
Db 290 ProThrProLysAlaGluThrThr-----LysGlyProAlaLeuThrThrProLys 307
Qy 204 TATCGACATCTACC-----TGCTCGACGACCCCGAACGAGCCCTTGCGC 248
Db 308 GluProThrProThrThrProLysGluProAlaSerThrThrProLysGluProThrPro 327
Qy 249 -----CCGCGCGCGCTTATCGAATCTGCCGAAACACACGCTCGATT 290
Db 328 ThrThrLeuLysSerAlaProThrThrProLysGluProAlaProThrThrThrLysSer 347
Qy 291 TCGAATACTACA-----TTTGGCGCAACGACATTT 320
Db 348 AlaProThrThrProLysGluProAlaProThrThrThrLysGluProAlaProThrThr 367
Qy 321 CCGGAGCGCTGCTGTCAACCTCATGTACCTTGGCGAGAACGCGCGGTGCGGTACGCC 380
Db 368 ProLysGluProAlaProThrThrThrLysGluProAla----- 380
```


DB:	11	Gaps:	29
US-10-685-990A-13 (1-1561) x US-11-169-232-2 (1-1404)			
Qy	144	CCAAACCTGTCCTCTCGAGCAACATCTCGAAATCCGGCACACCCCTCATAAACAACGGGC	203
Db	331	ProThrProLysAlaGluThrThr-----LysGlyProAlaLeuThrThrProLys	348
Qy	204	TATCCGACATCTACC-----TGCTCGACGACCCGCCAGAAAGTCCCTTCGCCG	248
Db	349	GluProThrProThrProLysGluProAlaSerThrThrProLysGluProThrPro	368
Qy	249	-----CCGCGCCCGCCTTATCGAATCTGCCGACACAGACGCTCGATT	290
Db	369	ThrThrIleLysSerAlaProThrThrProLysGluProAlaProThrThrThrLysSer	388
Qy	291	TGCAATACTACA-----TTTGGCGCAACGACATTT	320
Db	389	AlaProThrThrProLysGluProAlaProThrThrThrLysGluProAlaProThrThr	408
Qy	321	CCGCGAGGCTGCTTCAACCTCATGTACCTTGCCGACAGAACGGCGTGGCGTACGCC	380
Db	409	ProLysGluProAlaProThrThrThrLysGluProAla-----	421
Qy	381	TGCTGTTGGACGACAACAACGCGCGGTGGACGATCTCTGCTCGCCCTCGACAGCC	440
Db	422	-----ProThrThrThrLysSerAlaProThrThrProLysGluProAlaProThr	438
Qy	441	ATCCCAATATCGAAGTGGCTGTTCAACCCCTTCGTCCTACGCAAAATGGCGGCACTCG	500
Db	439	ThrPro-----LysLysProAlaProThrThrProLysGluProAlaProThrThrProLys	457
Qy	501	GCTACCTGACCGAGCTTCCCGCGCTCAACGCGCATGCACAACAAATCTCTTACCG--	557
Db	458	GluProThrProThrThrProLysGluProAlaProThrThrLysGluProAlaProThr	477
Qy	558	-----CCGACAAC-----	566
Db	478	ThrProLysGluProAlaProThrAlaProLysLysProAlaProThrThrProLysGlu	497
Qy	567	---CGGCCACATACTCGGGACGCAATATCGCGGAGCAATACTTCAAAGTCGGTAGG	623
Db	498	ProAlaProThrThrProLysGluProAlaProThrThrThrLysGluProSerProThr	517
Qy	624	ACACCGTT-----TGCCGACCTGGACATCTCGCCACCGGACGGTGGCGCGAAG	677
Db	518	ThrProLysGluProAlaProThrThrThrLysSerAlaProThrThrThrLysGluPro	537
Qy	678	TATCGACGACTTCGACCGTACTGGGCAAGCCATTCCGCCCAACAGCCACGCGCAGTCA	737
Db	538	AlaProThrThrThrLysSerAla-----ProThrThrProLysGluPro	552
Qy	738	TCCGACGGGCAACATCGGCAGGGTCTTCAAGCACTCGGNATCAACAGCAAGAAATCCA	797
Db	553	SerProThrThrThrLys-----GluProAlaProThrThrProLysGluPro	568
Qy	798	-----GACACGCGCTCCTCGCTACCGCGAAACCGTCGAACGTGCGCCCTCTACC	848
Db	569	AlaProThrThrProLysLysProAlaProThrThrThrProLysGluProAlaProThrThr	588
Qy	849	AAAAAATACAGACGGGACGATCGACTGGCAGAGCGTCCAAACCCGCTGATCAGCG---	905
Db	589	ProLysGluProAlaProThrThrThr-----LysLysProAlaProThrAlaPro	605
Qy	906	-----ACACCCCTGCAAAAGGACTCGACCGGACGCGCGCAACACCGC	947
Db	606	LysGluProAlaProThrThrPro-----LysGluThrAlaProThrThrProLysLys	623
Qy	948	CGATTGCGGGAGGCTCGAAGCGCTCAACAGCGCGGAAAAAGCGTCTATCTGGTTT	1007
Db	624	LeuThrPro-----ThrThrProGluLysLeuAla-----	633
Qy	1008	CACCTATTTTCGTCCTACAAAATCCGGCACAGACGCACTGGCAAACTGGTGCGAGACG	1067

ATTORNEY/AGENT INFORMATION:
NAME: Cserr, Luann
REGISTRATION NUMBER: 31,822
REFERENCE/DOCKET NUMBER: GI 5190
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 876-1170
TELEFAX: (617) 876-5851
INFORMATION FOR SEQ ID NO: 62:
SEQUENCE CHARACTERISTICS:
LENGTH: 1404 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 62:
US-11-169-232-62

Alignment Scores:
Pred. No.: 5,728-09 Length: 1404
Score: 247.50 Matches: 135
Percent Similarity: 35.2% Conservative: 44
Best Local Similarity: 26.5% Mismatches: 205
Query Match: 8.7% Indels: 125
DB: 11 Gaps: 29

US-10-665-990A-13 (1-1561) x US-11-169-232-62 (1-1404)

```
QY 144 CCAAACTGTCTCGACGACATCCTGCAAAATCGGCGACACCCCTCATACACGGGC 203
DB 331 ProThrProLysAlaGluThrThr-----LysGlyProAlaLeuThrProLys 348
QY 204 TATCCGACATCTACC-----TGCTCGACGACCCCGGCGAGCCCTTGCCG 248
DB 349 GluProThrProThrProLysGluProAlaSerThrThrProLysGluProThrPro 368
QY 249 -----CCGCGCGCGCTTATCGAATCTGCGGACACAGCCTCGATT 290
DB 369 ThrThrLeuLysSerAlaProThrThrProLysGluProAlaProThrThrLysSer 388
QY 291 TGCAATACTACA-----TTTGGCGCAACGACATTT 320
DB 389 AlaProThrProLysGluProAlaProThrThrLysGluProAlaProThrThr 408
QY 321 CCGCGAGGCTGTGTCAACCTCATGTACTTGGCGCAGAGACGGCGGTGCGGTACGCC 380
DB 409 ProLysGluProAlaProThrThrThrLysGluProAla----- 421
QY 381 TGCTGTGGACGACGACGACGCGCGGTGGAGCATCTCTGCTGCTGCTCGCCCTCGACAGCC 440
DB 422 -----ProThrThrThrLysSerAlaProThrThrProLysGluProAlaProThr 438
QY 441 ATCCCAATATCGAAGTGGCGCTGTTCACCCCTTCTGCTACGCAATGGCGCGCACTCG 500
DB 439 ThrPro-----LysLysProAlaProThrThrProLysGluProAlaProThrProLys 457
QY 501 GCTACCTGACGAGCTTCCCGCGCTCAACCGCGCATGCAACAATAATCTTTACCG--- 557
DB 458 GluProThrProThrProLysGluProAlaProThrThrLysGluProAlaProThr 477
QY 558 -----CCGCAACCC----- 566
DB 478 ThrProLysGluProAlaProThrAlaProLysLysProAlaProThrThrProLysGlu 497
QY 567 ---GGCCACCATACTCGGCGGCGCAATATCGGCGAGATACCTTCAAGTGGTGGAGG 623
DB 498 ProAlaProThrProLysGluProAlaProThrThrThrLysGluProSerProThr 517
QY 624 ACACCGTTT-----TCGCCGACCTGGACATCTCGCCACCGCGCGGTCTCGCGCAAG 677
DB 518 ThrProLysGluProAlaProThrThrThrLysSerAlaProThrThrThrLysGluPro 537
QY 678 TATCGCAGGACTTCGCGCTACTGGGCGAGCCATTCGCGCCACACGCGCGCATCA 737
```

```
DB 538 AlaProThrThrThrLysSerAla-----ProThrThrProLysGluPro 552
QY 738 TCCGCGAGCGCAACATCGGCAAGGTCTTCAAGCACTCGGATACACGAGAAACATCA 797
DB 553 SerProThrThrThrLys-----GluProAlaProThrThrProLysGluPro 568
QY 798 -----GACACGCGCTCTCGCTACCGCGAAACCGTCAACAGTCGCCCTCTAC 848
DB 569 AlaProThrThrProLysLysProAlaProThrThrProLysGluProAlaProThrThr 588
QY 849 AAAAAATACAGCGGCGCATCGACTGGCAGAGCGTCCAAACCCGCTGATCAGCG--- 905
DB 589 ProLysGluProAlaProThrThr-----LysLysProAlaProThrAlaPro 605
QY 906 -----ACACCTCTGCAAAAGGACTCGACGCGGACCGCGCAACCGC 947
DB 606 LysGluProAlaProThrThrPro-----LysGluThrAlaProThrThrProLysLys 623
QY 948 CGATTGCGGAGGCTGCAAGACGCGTCAACACGCGGCAAAAGCGTCTATCTGTTT 1007
DB 624 LeuThrPro-----ThrThrProGluLysLeuAla----- 633
QY 1008 CACCTATTTCCTTACAAAATCCGCGACAGCGCACTGGCAAACTGGTGCAGGAGC 1067
DB 634 -----ProThrThrProGluLysProAlaProThrThrProGlu-----GluLeu 648
QY 1068 GCATAGAGTTTACCGTCTGACCAACTCGCTACAGGCGACCGAGTTCGCG--- 1118
DB 649 AlaProThrThrProGluLysProThr-----ProThrThrProGluLysPro 664
QY 1119 ---CCGTCCATTCGCGGTACGTCAATACCGAAACCGCTCTCAAGCGCGCATCAAC 1175
DB 665 AlaProThrThrProLysAlaAla-----AlaProLysThrProLysGluProAla----- 681
QY 1176 TCTACGAGCTGCAACCAACGATGCGTCCCGCCACAAAGAGCAAGGCGTGCAGCGCA 1235
DB 682 -----ProThrThrProLysGluProAlaProThrThr-----ProLys 694
QY 1236 GCTCCGTAACGAGCTGCATGCCAAACCTTCATTGTGGAGGCGCAACGATCTTCATCG 1295
DB 695 GluProAlaProThr-----ThrProLysGluThrAlaProThrThrProLysGlyThrAla 713
QY 1296 GCTCATTAACCTCGACCCCGTTCGCGACGCGTCAATACCGAAATGGCGCTCGTCATCG 1355
DB 714 ProThrThrLeuLysGluProAlaPro-----ThrThrProLys-----Lys 727
QY 1356 AAGCCCAAAATCGCAGAACAGATGAGCGCACCTCGCGATACCAACCCGAATACG 1415
DB 728 ProAlaProLysGlu-----LeuAlaProThrThrThrLysGluProThrSer 743
QY 1416 CCTACCGGTTACCTCGACAAACACACACCGCTGCAATGGCAGATCCCG---CCAGCC 1472
DB 744 ThrThrSerAspLysProAlaProThrThrPro-----LysGlyThrAlaProThrThrPro 762
QY 1473 GAAAAACCTACCCCGCAAGCCCGAAG 1499
DB 763 LysGluProAlaProThrThrProLys 771
```

RESULT 24

US-10-467-657-5700
; Sequence 5700, Application US/10467657
; Publication No. US20050260581A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON SpA
; APPLICANT: FONTANA Maria Rita
; APPLICANT: PIZZA Mariagrazia
; APPLICANT: MASNIGNANI Vega
; APPLICANT: MONACI Elisabetta
; TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/467,657
; CURRENT FILING DATE: 2003-08-11
; PRIOR APPLICATION NUMBER: GB-0103424.8

```
; PRIOR FILING DATE: 2001-02-12
; NUMBER OF SEQ ID NOS: 9218
; SOFTWARE: SeqWin99, version 1.04
; SEQ ID NO 5700
; LENGTH: 430
; TYPE: PRT
; ORGANISM: Neisseria gonorrhoeae
US-10-467-657-5700

Alignment Scores:
Pred. No.:          9e-09          Length:          430
Score:             243.50         Matches:         123
Percent Similarity: 33.7%         Conservative:    37
Best Local Similarity: 25.9%      Mismatches:     166
Query Match:       8.5%          Indels:         149
DB:                9            Gaps:           24

US-10-665-990A-13 (1-1561) x US-10-467-657-5700 (1-430)
QY 310 CAACGACATTTCCGCGAGGCTGCTGTTCAACCTCATGTACCTTGGCGGAGAACGCGGCGT 369
Db 21 GluGlnArgIleArgAsnAlaValGlyGlySerIleProPheSerGlyThrAsnThr 40
QY 370 GCGCGTACGCGTCTGTTGGACGCAACACACGCGCGGTTGGACGATCTCTCTCGC 429
Db 41 LysHisArgProSerLeuGlyLysProValLeuPheHisIleSerAlaSerAspGly 59
QY 430 CCTCGACAGCCATCCCAATATCGAAGTGGCGCTGTTCAA---CCCTTTGCTCT---ACG 483
Db 60 -----IleArgTyrGluLeuLeuAlaThrGlnThrProPheArgProLeuAla 75
QY 484 CAAATGGCGCGCACTCGGCTACTGACCGACTTCCCGCGCTCAACCGCGGATGACAA 543
Db 76 GlnLeuLeuArgThr-----LeuProProPro---ProProHisThrCys 89
QY 544 CAAATCTTTACCGCGCAACCG---CGCACCATCT-----CGG 579
Db 90 ArgAlaAlaArgArgAsnArgProValArgHisArgThrArgProAlaThrProPro 109
QY 580 -----CGCGGAGCGCAATAT-----CGG 597
Db 110 ThrArgArgMetAspArgAspArgLeuArgArgProArgHisAlaProValProArg 129
QY 598 CGACGAATACTTCAA-----ATCGGTGAGGACACCGTTTTCGCCGA--- 639
Db 130 ArgAspLeuLeuGlnGlyGlyGlyThrTyrAlaAlaGlyTyrHisArgAlaGlyArgGly 149
QY 640 -----CCTGGACATCTCGC 654
Db 150 PheGlyArgPheMetAlaGluProAlaLeuPheProArgGlnProProLeuLeuProAsp 169
QY 655 CACCGCGACGTCGTCGGCGAAGT-----ATCGCAGCATCTCGACCG 696
Db 170 HisArgHisGlyLysArgThrGlyArgLeuGlyGlyArgGlnLysArgLeuArgPro 189
QY 697 CTACTGGGAAGCATTCGCCCCACACCCCGCGCATCATCCGAGCGGCAACATCGG 756
Db 190 TyrAlaGlyGly-----AlaAspAspValHisAlaHisArgArgGlnArgGlnArgMet 207
QY 757 CAAGGTCTTCAAGCACTCGGATACAGCAGCAACGAAACATCCAGACGCGCTCTCGGCTA 816
Db 208 AlaArg-----GlnArgProAspAlaArgAspGluArgProHisArg 221
QY 817 CGCGGAAC-----CGTGAACAGTCCGCCCTCTTACCAAAAAAATAACAGCGGAGC 867
Db 222 ArgArgHisArgHisCysArgArgGlnThrAlaAlaGluIleHisThrAspValAla 241
QY 868 CATCGACTGGCAGAGGTCCTCAAAACCGCTGATCAGCGACCCCTCGAAAAGGACTCGA 927
Db 242 Phe-----HisAlaCysArg-----Gln 247
QY 928 CCGCGACCGCGCAACCGCGGATTCGCGGAGGCTGCAAGACGCGCTCAACAGCCCGA 987
Db ----- 987

US-10-467-657-7448
QY 248 ProGlyArgLeuGlnGlnAsnAspCysArgAsnGlnGlnArgGlnAlaTyrAspAlaArg 267
QY 988 AAAAAAGCGTCTATCTGTTTCCACCTATTTCCGTCCTACAAAATCCGCGCAGACGCACT 1047
Db 267 ----- 267
QY 1048 GGCAAAACCTGTCAGGACGCATAGAGTTACCGTCTCTGACCACTCGCTACAGGCGAC 1107
Db 268 -----ThrPheGlyAlaGluTyrGlyGlnAsnAlaProAsnGlnArgThrHisGlyGln 285
QY 1108 CGACGTTGCCCGCTCCATTCGGCTCAATACCGATAACGAAACCGCTGCTCAAGCCGG 1167
Db 286 Lys-----ProGlnProProArg---ArgHisIleGlyArgLysProHisGlnProLeu 302
QY 1168 CATCAAACTCTACGAGCTGCAACCAACCATGCGCTGCCCGCCCAAAAGACAAAGGCGCT 1227
Db 303 HisAspGlySerHisAlaAlaArgProProGlnAsnArgGlnHisArgAlaAlaPro 322
QY 1228 GACCGGAGCTCCGTAACCGCTGCATGCCAAAACCTTCATTGTGGAGCGCAACGCAT 1287
Db 323 AspHis-----ArgArgGlnAlaAlaIleSer-----GlnThrGln 334
QY 1288 CTTTCATCGGCTCAITTCACCTCGACCCCGCTCCGCGACGGCTCAATACCGAAATGGCGCT 1347
Db 335 ArgGlnArgAsnProAlaAlaArgProProLeuHisThrAlaProAsnArgProAla--- 353
QY 1348 CGTCATCAAAAGCCCCAAAATTCGAGAACACATGGAGGCGACCCCTCGCGCATACACACC 1407
Db 354 ThrAsnArgArgProHisGlnArgGlnThrArgProProHisProHisArgHisArgHis 373
QY 1408 CGAATACGCTCTA----- 1419
Db 374 GlnProArgThrGlySerProArgThrProProLeuProMetAlaGlyLeuProLeu 393
QY 1420 -----CGCGGTTACCTTCGACCAACCAACCTACCGGAA 1488
Db 394 AlaGlnHisGlnTyrAlaSerGlyAsnPheArgProArgHisProAlaAlaThrHisPro 413
QY 1447 CTGCAATGGCAGCA---TCCCGCGCACCCGAAACCTACCGGAA 1488
Db 414 ProGlnMetAlaGlyCysProArgThrProThrProAlaProLys 428

RESULT 25
US-10-467-657-7448
; Sequence 7448, Application US/10467657
; Publication No. US20050260581A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON Spa
; APPLICANT: FONTANA Maria Rita
; APPLICANT: PIZZA Mariagrazia
; APPLICANT: MASIGNANI Vega
; APPLICANT: MONACI Elisabetta
; TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/467,657
; CURRENT FILING DATE: 2003-08-11
; PRIOR APPLICATION NUMBER: GB-0103424.8
; PRIOR FILING DATE: 2001-02-12
; NUMBER OF SEQ ID NOS: 9218
; SOFTWARE: SeqWin99, version 1.04
; SEQ ID NO 7448
; LENGTH: 430
; TYPE: PRT
; ORGANISM: Neisseria gonorrhoeae
US-10-467-657-7448

Alignment Scores:
Pred. No.:          9e-09          Length:          430
Score:             243.50         Matches:         123
Percent Similarity: 33.7%         Conservative:    37
Best Local Similarity: 25.9%      Mismatches:     166
Query Match:       8.5%          Indels:         149
DB:                9            Gaps:           24
```

Db 303 HisAspGlySerHisAlaAlaAArgProGlnAsnArgGlnHisHisArgAlaAlaPro 322
 QY 1228 GACGGCAGCTCCGTAACACGACCTCGCATGCCAAAACCTTCATTGTGGACGGCAAAACGCAT 1287
 Db 323 Asphis-----ArgArgGlnAlaAlaSer-----GlnThrGln 334
 QY 1288 CTTTCATCGGCTCATTAACCTCGACCCCGTTCGCGACGGCTCAATACCGAAATGGCGCT 1347
 Db 335 ArgGlnArgAsnProAlaAlaAArgProProLeuHisThrAlaProAsnArgProAla-- 353
 QY 1348 CGTTCATCGAAGCCCAAAATCCGAGAACAGATGGAGCGCACCTCCCGATACCACACC 1407
 Db 354 ThrAsnArgArgProHisGlnArgGlnThrArgProProHisProHisArgHisArgHis 373
 QY 1408 CGAATACGCCTA----- 1419
 Db 374 GlnProArgThrGlySerProArgArgThrProProLeuProMetAlaGlyLeuProLeu 393
 QY 1420 -----CCGGTTACCTCGACAAACACACACCG 1446
 Db 394 AlaGlnHisGlnTyrAlaSerGlyAsnPheArgProArgHisProAlaAlaThrHisPro 413
 QY 1447 CCTCAATGCCAGA---TCCCGCCACCCGAAAAACCTACCCGAA 1488
 Db 414 ProGlnMetAlaGlyCysProArgThrProThrProAlaProLys 428

RESULT 26
 US-11-079-463-7133
 ; Sequence 7133, Application US/11079463
 ; Publication No. US20060073161A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Gary L. Breton
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO BACTERIA
 ; FILE REFERENCE: PATH00-03DIV2
 ; CURRENT APPLICATION NUMBER: US/11/079,463
 ; PRIOR FILING DATE: 2005-03-14
 ; PRIOR APPLICATION NUMBER: US 60/128,705
 ; PRIOR FILING DATE: 1999-04-09
 ; PRIOR APPLICATION NUMBER: US 09/540,209
 ; PRIOR FILING DATE: 2000-04-04
 ; NUMBER OF SEQ ID NOS: 10444
 ; SEQ ID NO 7133
 ; LENGTH: 503
 ; TYPE: PRT
 ; ORGANISM: B. fragilis
 US-11-079-463-7133

Alignment Scores:
 Pred. No.: 9,19e-09 Length: 503
 Score: 243.50 Matches: 112
 Percent Similarity: 37.0% Conservative: 69
 Best Local Similarity: 22.9% Mismatches: 184
 Query Match: 8.5% Indels: 125
 DB: 11 Gaps: 16

US-10-665-990A-13 (1-1561) x US-11-079-463-7133 (1-503)

QY 32 TCCCGATGAACACGACGCTCATTTCCCTTTATGCCTCTCTT----- 78
 Db 61 AsnProValIysThrMetAlaTrpIleLeuLeuMetPheLeuProValValGlyLeu 80
 QY 79 CTGTTCATGTTCTTCATGTTGCCCCCTACCTGGAAGAACGGAC----- 120
 Db 81 ValPheTyrrPhePhe-----GlyArgSerGlnArgArgGluIysIleIle 96
 QY 121 GGAAGACCGCTCAPTTTCAATACTTCCAAACCTGTCCTCGGACCAACATCCTGCA----- 174
 Db 97 GlyIysIysSerTyrrAspArgLeuLeuLysLysProMetAlaGluTyrrLeuAlaGlnAsn 116
 QY 175 -----AATCG 180


```
Qy 181 GCACACCCTCATACAAACGGGCTATCCGACATCTACCTGCTCGACGACCCCGCAGAGC 240
Db   |||
Qy 137 AlaPhe-ProPheGluGlyAsnArgValAspIleTyr-----ThrGlyGlyTyrSerLy 154
Db   |||
Qy 241 CTTGGCGCGCGCGCCCTTATCGAATCTCGCGAACACACACGCTCGATTTGGCAACTA 300
Db   |||
Qy 154 sLeuGlnAlaLeuLeuArgGluLeuGlnIleAlaArgLeuHisIleHisMetGluTyrTy 174
Db   |||
Qy 301 CATTGGCGCAACGACATTTCCGGCAGGCTGCTGTTCACCTCATCTACCTTGGCGCAGA 360
Db   |||
Qy 174 rIlePheGluAspAspProValGlyArgLeuValArgAspValLeuIleGluLysAlaAr 194
Db   |||
Qy 361 ACGCGGCTGCGGTACGCTGCTGTTGGAGACAAACACGCGCGGTTGGAGATCT 420
Db   |||
Qy 194 gGluGlyValGluValArgValIleTyrAspAspValGlyCysTrpHisValProHisAr 214
Db   |||
Qy 421 CCTGCTGCGCTCGACAGCCATCCCAATATCGAAGTGGCTGTTCACCCCTTCGTCT 480
Db   |||
Qy 214 gPhePheGluGluMetArgAspAlaGlyIleGluValArgSerPhe----- 229
Db   |||
Qy 481 ACGCAATATGGCGGCACTCGGCTACCTGACGAGCTTCCCC-----CGCCTCAA 528
Db   |||
Qy 230 -----LeuLysValArgPheProLeuPheThrSerLysValAs 242
Db   |||
Qy 529 CGCGCGCATGCACAAACAAATCTTTACCGCGACAAACCGCGCCACCATCTCGCGGAGC 588
Db   |||
Qy 242 nTyrArgAsnHisArgLysIleValIleAspGlyArgIleGlyPheIleGlyGlyMe 262
Db   |||
Qy 589 CAATATCGGCGACGAATATCTCAAGTCGGTGAGGACACCGCTTTTCGCGACCTGGACAT 648
Db   |||
Qy 262 tAsnLeuAlaGluArgTyrMetArggLyPheSerTrpGlyIleTrpArgAspThrHisil 282
Db   |||
Qy 649 CCTGCCACCGCGAGCGTGTGGCGGAAGTATCGCACGACTTC-----GACCGCTACTG 702
Db   |||
Qy 282 eLeuLeuGluGlyLysAlaValHisGlyLeuGlnThrAlaPheLeuLeuAspTrpTyrPh 302
Db   |||
Qy 703 GCGAAGCCATTCGCGCCCAACAGCGCACGCGGATCATCCGCGAGCGCAACATCGCGAAGG 762
Db   |||
Qy 302 eValAspArgThrLeuIleThrAlaSerArgTyrPheProLys----- 316
Db   |||
Qy 763 TCTTCAAGCACTCGGATACACAGCAAGCAACATCCAGACACGCGCTCTCGCTACCGGGA 822
Db   |||
Qy 317 -IleGluAlaTyrGly-----AsnSerLeuVal-----G1 326
Db   |||
Qy 823 AACGTCGAACAGTGCCTC-----TACCAAAAAATACAGACGGGACGCATCGACTG 876
Db   |||
Qy 336 nIleValThrSerGluProIleGlyProTrpLysGluIleMetGlnGly----- 342
Db   |||
Qy 877 GCAGAGCGTCNAACCCCGCTGATCAGCGACACCCCTCCAAAGACTCGACCGGACCG 936
Db   |||
Qy 342 ----- 342
Qy 937 CCGCAACCGCGATTTGCGGAGGCTGCAAGACGCGCTCAAAACAGCCCGCAAAAAGCGT 996
Db   |||
Qy 343 -----LeuThrValAlaIleSerGlyAlaLysLysTyrPh 354
Db   |||
Qy 997 CTATCTGTTTACACCTATTTCCCTACAAAATCCGGCACAGACGCACTGGCAAAACT 1056
Db   |||
Qy 354 eTyrMetGlnThrProTyrPheLeuProThrGluGlnIleLeuGlyAlaMetGlnThrAl 374
Db   |||
Qy 1057 GGTGACGACGCGATAGAGTTACCGTCTGACCACTCGCTACAGGCGACGAGTTGC 1116
Db   |||
Qy 374 aAlaLeuAlaGlyValAlaAspIleArgLeuMetLeuPro-----GluHisAlaAspAsnAr 392
Db   |||
Qy 1117 CGCGTCTCATTCGGCTACGTCAAAATACGGAACCGCTGTCTCAAGCGCGCATCAAACT 1176
Db   |||
Qy 392 gValThrHisLeuGlySerCysSerTyrLeuAlaAspValLeuArgAlaGlyValLysVa 412
Db   |||
Qy 1177 CTACGAGTGCACCAACCATTCGCGTCCCGGCCCAAAAAAGACCAAGCGCTGACCGGAG 1236
Db   |||
Qy 412 lTyrPheTyrLys-----LysGlyPhe----- 419
```

```
Qy 1237 CTCGTAACACGAGCTGATGCCAAAACCTTCATTGTCGACGGCAACGCATCTTCATCGG 1296
Db   |||
Qy 420 -----LeuHisSerLysLeuMetValSerAspMetLeuSerThrValGI 435
Db   |||
Qy 1297 CTCAATCAACCTCGACCCCGTTCCGACGCGCTCAATACCGAAATGGCGCTGTCATCGA 1356
Db   |||
Qy 435 ySerThrAsnLeuAspPheArgSerPheGluHisAsnPheGluValAsnAlaPheMetTy 455
Db   |||
Qy 1357 AAGCCCCAAAATCGCAGAACAGATG 1381
Db   |||
Qy 455 rAspMetGluThrAlaLeuGlnMet 463
Db   |||

RESULT 27
US-10-873-528-109
; Sequence 109, Application US/10873528
; Publication No. US20050276814A1
; GENERAL INFORMATION:
; APPLICANT: Microbial Technics Limited
; APPLICANT: Gilbert, Christophe FG
; APPLICANT: Hansbro, Phillip M
; TITLE OF INVENTION: Proteins
; FILE REFERENCE: PMC/P21129WO
; CURRENT APPLICATION NUMBER: US/10/873,528
; PRIOR FILING DATE: 2004-06-23
; PRIOR FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: GB 9816337.1
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: US 60/125164
; PRIOR FILING DATE: 1999-03-19
; NUMBER OF SEQ ID NOS: 388
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 109
; TYPE: PRT
; LENGTH: 1236
; ORGANISM: Streptococcus pneumoniae
US-10-873-528-109

Alignment Scores:
Pred. No.: 1,04e-08 Length: 1236
Score: 243.50 Matches: 124
Percent Similarity: 42.7% Conservative: 94
Best Local Similarity: 24.3% Mismatches: 219
Query Match: 8.5% Indels: 74
DB: Gaps: 16

US-10-665-990A-13 (1-1561) x US-10-873-528-109 (1-1236)

Qy 126 GCGGTCAATTTCAATACCTTCCAAACCTGCTCTCTCGGCAACATCTCGAAAATCGGCACA 185
Db   |||
Qy 551 AlaSerAlaSerIleSerAlaSerGluSerAlaSerThrSerAlaSerGluSerAlaSer 570
Db   |||
Qy 186 CCCCTCAACAAACGCGCTATCCGACATCT-----ACCTGCTCGACGCCCGCAG 236
Db   |||
Qy 571 ThrSerThrSerAlaSerAlaSerThrSerAlaSerGluSerAlaSerThrSerAlaSer 590
Db   |||
Qy 237 AAGCCCTTGGCGCGCGCGCCCTTATCGAATCTCGGAACACAGCGCTCGATTCGCAAT 296
Db   |||
Qy 591 AlaSerAlaSerThrSerAlaSerAlaSerAlaSerAlaSerThrSerAlaSerAlaSer 610
Db   |||
Qy 297 ACTACATTTGGCGCAACGACATTTCCGCGAGGCTGTCTTCAACCTCATGTACCTTGCGCG 356
Db   |||
Qy 611 ThrSerAlaSerAlaSerThrSerAlaSerGluSerAlaSerThrSerAlaSerAlaSer 630
Db   |||
Qy 357 CAGAACGCGCGCTGCGCTACGCCCTGCTGTTGGACGACAAACACGCGCGGTTGGAGC 416
Db   |||
Qy 631 AlaSerThrSerAlaSerAlaSerAlaSerThrSerAlaSerAlaSerAlaSerThrSer 650
Db   |||
Qy 417 ATCTCTGCTGCGCTCGACGAGCCATCCCATATCG---AAGTGGCGCTGTTCACCCCT 473
Db   |||
Qy 651 AlaSerAlaSerAlaSerThrSerAlaSerAlaSerAlaSerThrSerAlaSerAlaSer 670
Db   |||
Qy 474 TCGTCTACGCAAAATGGCGCGCACTCGGCTACCTGACCGAGCTTCCCCCGCTCAACCGCC 533
```

```
Db 671 AlaserThrSerAlaserAlaserAlaserThrSerAlaserGluSerAlaserThrSer 690
      ::::: ::::: ||| ::::: ||| ::::: ||| ::::: ||| ::::: ||| ::::: |||
Qy 534 GCATGCACAAACATCTTTACCGCGGACCAACCGCGCCACCATCTCGCGGACCAATA 593
      ::::: ||| ::::: ||| ::::: ||| ::::: ||| ::::: ||| ::::: |||
Db 691 AlaserAlaserThrSerAlaserAlaserAlaserThrSerAlaserAlaser 710
      ::::: ||| ::::: ||| ::::: ||| ::::: ||| ::::: ||| ::::: |||
Qy 594 TCGCGACGAATCTTCAAGTCGGTGAGGACACCGTTT----- 632
      ::::: ||| ::::: ||| ::::: ||| ::::: ||| ::::: ||| ::::: |||
Db 711 AlaserThrSerAlaserGluSerAlaserAlaserThrSerAlaserAlaserThrSer 730
      ::::: ||| ::::: ||| ::::: ||| ::::: ||| ::::: ||| ::::: |||
Qy 633 -----TGCAGACTGACATCTCTCGCCACCGGACGCGTCTCGGGAAGTATCGC 683
      ||| ::::: ||| ::::: ||| ::::: ||| ::::: ||| ::::: |||
Db 731 AlaserGluSerAlaserThrSerAlaserAlaserAlaserThrSerAlaserAlaser 750
      ||| ::::: ||| ::::: ||| ::::: ||| ::::: ||| ::::: |||
Qy 684 ACGACTTCAGCGCTACTGGGCAAGCCATTCGCGCCACCAACGCCAGCGCATCTCCGCA 743
      ::::: ||| ::::: ||| ::::: ||| ::::: ||| ::::: ||| ::::: |||
Db 751 AlaserThrSerAlaserGlySerAlaserThrSerThrSerAlaserAlaserThrSer 770
      ::::: ||| ::::: ||| ::::: ||| ::::: ||| ::::: ||| ::::: |||
Qy 744 CGCGCAACATCGGCAAGGGTC-----TTCAAGCACTCGGATACACGACGAAACAT 794
      ||| ::::: ||| ::::: ||| ::::: ||| ::::: ||| ::::: |||
Db 771 AlaserAlaserAlaserThrSerAlaserAlaserAlaserIleSerAlaserGluSer 790
      ||| ::::: ||| ::::: ||| ::::: ||| ::::: ||| ::::: |||
Qy 795 CCAGACAGCGCTCTCGGCTACCGCGAAACGGTCGAACAGTCGCCCTCTACCAAAAAA 854
      ||| ::::: ||| ::::: ||| ::::: ||| ::::: ||| ::::: |||
Db 791 AlaserThrSerAlaserGluSerAlaserThrSerThrSerAlaserAlaserThrSer 810
      ||| ::::: ||| ::::: ||| ::::: ||| ::::: ||| ::::: |||
Qy 855 TACAGCGGACGCTGCGGACGCGTCCAAACCGCGCTGATCAGCGCACCCCTCG 914
      ||| ::::: ||| ::::: ||| ::::: ||| ::::: ||| ::::: |||
Db 811 AlaserGluSerAlaserThrSerAlaserAlaserAlaserThrSerAlaserAla--- 829
      ||| ::::: ||| ::::: ||| ::::: ||| ::::: ||| ::::: |||
Qy 915 CAAAAGGACTCGACCGCACCGCGCAACCGCGGATTCGCGGAGGTCGAAGACGCGC 974
      ||| ::::: ||| ::::: ||| ::::: ||| ::::: ||| ::::: |||
Db 830 -----SerAlaserThrSerAla-SerAlaserAlaArgGln--ValArgArgPr 845
      ||| ::::: ||| ::::: ||| ::::: ||| ::::: ||| ::::: |||
Qy 975 TCAACAGCCCGAAAAAGCGCTCTATCTGGTTTACCCTATTCTGCTCCTACAAATCCG 1034
      ||| ::::: ||| ::::: ||| ::::: ||| ::::: ||| ::::: |||
Db 845 oGlnProValHis-----LeuAsnArgHisGlnProValAr 857
      ||| ::::: ||| ::::: ||| ::::: ||| ::::: ||| ::::: |||
Qy 1035 GCACAGAGCACTGGCAAACTGGTGCAGGACGATAGAGCTTACCGTCTCGACCACT 1094
      ||| ::::: ||| ::::: ||| ::::: ||| ::::: ||| ::::: |||
Db 857 gGlnProGlnValLeuValHisGlnLeuGlnHisGlnArgValHisArgLeuGlnHi 877
      ||| ::::: ||| ::::: ||| ::::: ||| ::::: ||| ::::: |||
Qy 1095 CGCTACAGGCGACGAGCTTCCCGCGCTCCATTCGG---CTAGCTCAAAATACCGAAAA 1151
      ||| ::::: ||| ::::: ||| ::::: ||| ::::: ||| ::::: |||
Db 877 sGlnProValProArgLeuGlnArgGlnProValArgGlnLeuGlnValPro----- 895
      ||| ::::: ||| ::::: ||| ::::: ||| ::::: ||| ::::: |||
Qy 1152 CGTGTCTAAAGCGGATCAAACTCTACGAGCTGCAACCCCAACCATGCGCTCCCGCCA 1211
      ||| ::::: ||| ::::: ||| ::::: ||| ::::: ||| ::::: |||
Db 896 -ValLeuGlnSerGlnHisGln-----GlnValLeuGlnProGlnHi 909
      ||| ::::: ||| ::::: ||| ::::: ||| ::::: ||| ::::: |||
Qy 1212 CAAAAGACAAAGGCTGACGCGGAGCTCGTAACAGCCTCATTCGCTCATGCCAAACCTTATTG 1271
      ||| ::::: ||| ::::: ||| ::::: ||| ::::: ||| ::::: |||
Db 909 s--ArgGlnValPro-----ArgLeuGlnAlaHisGlnHisLeuAsnGl 924
      ||| ::::: ||| ::::: ||| ::::: ||| ::::: ||| ::::: |||
Qy 1272 TGGACGGCAACCGCATCTTCATCGGCTCATTCACCTCGACCCCGTTCCGACGGCTCA 1331
      ||| ::::: ||| ::::: ||| ::::: ||| ::::: ||| ::::: |||
Db 924 nArgGlnAlaProGlnLeuGlnValProValArgGlnProGlnArg----- 941
      ||| ::::: ||| ::::: ||| ::::: ||| ::::: ||| ::::: |||
Qy 1332 ATACCGAAATGGGGCTGTCATCGAAAGCCCCCAAAATCGCGAGAACAGATGGAGCGCACCC 1391
      ||| ::::: ||| ::::: ||| ::::: ||| ::::: ||| ::::: |||
Db 942 -----ArgGlnValArgGlnProGlnGlnValLeuValHisGlnLeuGlnHi 957
      ||| ::::: ||| ::::: ||| ::::: ||| ::::: ||| ::::: |||
Qy 1392 TCGCGATACCAACCCGAAATACCGCTACCG----- 1422
      ||| ::::: ||| ::::: ||| ::::: ||| ::::: ||| ::::: |||
Db 957 sGlnArgValHisArgLeuArgArgGlnProValHisGlnSerGlnGlnValProValAr 977
      ||| ::::: ||| ::::: ||| ::::: ||| ::::: ||| ::::: |||
Qy 1423 -----GTTTACCCTCGACAAACAAACACCGCTCGCA-----ATGGCAGC 1460
      ||| ::::: ||| ::::: ||| ::::: ||| ::::: ||| ::::: |||
Db 977 gGlnLeuProHisGlnGlnValProArgLeuGlnGlnAlaProValArgArgLeuGlnGl 997
      ||| ::::: ||| ::::: ||| ::::: ||| ::::: ||| ::::: |||
Qy 1461 ATCCCGCCACCGAAAAACCTTACCGGAACGAAACCCGA-----AGCCAAAC 1505
      ||| ::::: ||| ::::: ||| ::::: ||| ::::: ||| ::::: |||
```

```
Db 997 nValLeuAlaProGlnProGlnProGlnProValArgGlnProGlnGlnValSerGlnAr 1017
Qy 1506 TTTGGAAACGATCGCCGCAAAATCTCTATC 1536
      ||| ::::: ||| ::::: ||| ::::: ||| ::::: ||| ::::: |||
Db 1017 gLeuAsnArgHisGlnArgValArgProLeu 1027
      ||| ::::: ||| ::::: ||| ::::: ||| ::::: ||| ::::: |||
RESULT 28
US-11-096-568A-14593
; Sequence 14593, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nickolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptide
; FILE OF INVENTION: Therby
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 14593
; LENGTH: 758
; TYPE: PRT
; ORGANISM: Zea mays subsp. mays
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(758)
; OTHER INFORMATION: Ceres Seq. ID no. 11049133
US-11-096-568A-14593
Alignment Scores:
Pred. No.: 1.32e-08 Length: 758
Score: 241.50 Matches: 150
Percent Similarity: 30.6% Conservative: 49
Best Local Similarity: 23.0% Mismatches: 187
Query Match: 8.5% Indels: 265
DB: 11 Gaps: 32
US-10-665-990A-13 (1-1561) x US-11-096-568A-14593 (1-758)
Qy 160 GGCAACATCTGCAATCCGGCACACCCCTCATAAACAGGGCTATCCGACATCTACT 219
      ||| ::::: ||| ::::: ||| ::::: ||| ::::: ||| ::::: |||
Db 83 GlyLeuHisPro-----ProArgArgProGlyArgArgArgLeuArgGlnArg 100
      ||| ::::: ||| ::::: ||| ::::: ||| ::::: ||| ::::: |||
Qy 220 GCTCGACGA-----CCCCCAGAAAGCCCTTCCGCGCGCGC 255
      ||| ::::: ||| ::::: ||| ::::: ||| ::::: ||| ::::: |||
Db 101 ValArgArgArgArgArgAlaArgArgGlyGlyProProArgGlnGlnArg 120
      ||| ::::: ||| ::::: ||| ::::: ||| ::::: ||| ::::: |||
Qy 256 CGCCCTTATCGAATCTCCGAACACACGCTCGATTGCAATCTACTATTTCGCGCAACGA 315
      ||| ::::: ||| ::::: ||| ::::: ||| ::::: ||| ::::: |||
Db 121 ValAspProArgLeuHisArgProArgArg-----HisLeu---ArgArg 135
      ||| ::::: ||| ::::: ||| ::::: ||| ::::: ||| ::::: |||
Qy 316 CATTTCCGCGGCTGCTGTTCACCTCATCTGTCGCGCAGAACCGCGCTGCGCT 375
      ||| ::::: ||| ::::: ||| ::::: ||| ::::: ||| ::::: |||
Db 136 HisHisArg-----LeuArgArgHisLeuProProHisGlnGlnArg 152
      ||| ::::: ||| ::::: ||| ::::: ||| ::::: ||| ::::: |||
Qy 376 ACGCTCTCTTGG-----CGACAA 396
      ||| ::::: ||| ::::: ||| ::::: ||| ::::: ||| ::::: |||
Db 153 AlaProGlyArgAlaAlaGlnAlaSerGlnArgArgAsnLeuArgHisArgGlnArg 172
      ||| ::::: ||| ::::: ||| ::::: ||| ::::: ||| ::::: |||
Qy 397 CAACACCGCGGTTGA----- 414
      ||| ::::: ||| ::::: ||| ::::: ||| ::::: ||| ::::: |||
Db 173 AlaHisAlaValGlyHisProArgGlyAspAlaGlyAlaHisGlnHisProPro 192
      ||| ::::: ||| ::::: ||| ::::: ||| ::::: ||| ::::: |||
Qy 415 -----CGATCTCTCTCGCTCGCCCTCGACAGCCATCCCAATAT 450
      ||| ::::: ||| ::::: ||| ::::: ||| ::::: ||| ::::: |||
Db 193 ProGlyLeuLeuArgHisProLeuArgAspProArgGlyHisHisGlnHisHis 212
      ||| ::::: ||| ::::: ||| ::::: ||| ::::: ||| ::::: |||
Qy 451 CGAAGTCGCGCTGTTCAACCCCTTCGCTACGCAAAATGGCGCGCACTCGCGTACTCT 507
      ||| ::::: ||| ::::: ||| ::::: ||| ::::: ||| ::::: |||
Db 213 ArgCysGlnProLeu---ProAlaAlaProGlyHisHisHisArgValGlyArgProGly 231
      ||| ::::: ||| ::::: ||| ::::: ||| ::::: ||| ::::: |||
Qy 508 -----GACCGACTTCCCGCGCTCAACCGCGCATGCACAAACAAATCTCTTAC 555
      ||| ::::: ||| ::::: ||| ::::: ||| ::::: ||| ::::: |||
Db 232 ProAlaLeuLeuHisArgArgProHisHisGlyProPro----- 244
      ||| ::::: ||| ::::: ||| ::::: ||| ::::: ||| ::::: |||
```

```
QY 556 CGCCGACACCGCGCCACCATCTCGCGGAGCAATATCGCGCAGCAATCTTCAAAGT 615
Db :||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
245 GlnArgAlaGlyArgHisArgArgArgGlyGlyArgArgGlyValGlnAsp 264
QY 616 CGGTAGGACACCGTTTTCGCG-----CGACCT---GGACATCTCGCCACCGG 660
Db :||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
265 Arg---ArgHisArgGlyArgLeuGlnAlaGlnProGlnGlnGlyProArgHisArg 283
QY 661 CAGCGTCGT-----CGCGCAAGTATCGCAGCACTTCGCGCG 696
Db :||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
284 GlnArgHisValArgGlyLeuArgAlaArgGlyHisArgAspValArgGlnArgPro 303
QY 697 -----CTACTGGCAAGCCATTCGCG----- 717
Db :||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
304 GlyArgProValGlyGlyProValArgLeuLeuArgGlyHisGluArgGlnAlaArg 323
QY 718 -----CCACA----- 723
Db :||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
324 ValHisGlyProProAspProGlnAlaGluAlaProProGlyValHisArgGlyArgGly 343
QY 724 CGCCACGCGCATCTCCG-----CAGCGGCAACATCGCAAGGTCTTCAAGCACTCGG 777
Db :||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
344 HisHisGlyAlaHisProGlyTrpGlnLeuLeuHisGluAlaGlyGlnGlyGluArg 363
QY 778 -----ATACAACGACGAAACATCCAGACACGCGCTCTCGCGCTACCGCAAC 825
Db :||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
364 AlaGlyProAlaAlaGluAlaGlnAlaGlyGlnValArgAlaPro----- 378
QY 826 CGTCGAACAGTCGCCCTCTACCAAAATACAGACGGGACGCATCGCTGCACAGCGT 885
Db :||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
379 -----HisValAlaAlaVal 383
QY 886 CCAAAACCGCGCTGATCAGCGACACCCC----- 912
Db :||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
384 AlaGlyProProAspArgGlyHisProArgArgHisGlnValHisArgAlaArgGlyGln 403
QY 913 TCGAAAGACTCGACCG-----CGACCGCGCAACCGCCGATTCGCGGAGGCT 963
Db :||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
404 LeuArgGluArgGlnProGlyHisArgArgProArg----- 416
QY 964 GCAACAGCGCTCAACAGCCGCAAAAGCGTCTATCTGTTTCAACCTATTTCTGTC 1023
Db :||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
417 -----GlnGlyAlaAlaArgGlnLeuProGlyHisPro----- 428
QY 1024 TACAAAATCCGCGACAGCGACTCGCAAAACTGGTGCAGGACGCATAGACTTACCGT 1083
Db :||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
429 -----HisArgArgValHisGlyGlnArgProProArgHisArgGlnHisArg 444
QY 1084 CTGACCAACTCGCTACAGGACCGACGTTCCCGCGCTCCATTCGCGCTACGTCAAATA 1143
Db :||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
445 GlnAlaHisValArgAlaValLeuArgAlaArgGlnArgValLeuGlnGlnArgAlaHis 464
QY 1144 CGCAAAACCGCTGCTCAAGCCGGCATCA----- 1173
Db :||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
465 LeuGlnProGlyArgGlnProGlnProGlnProGlyLeuArgLeuGlnGlyHisArgAsp 484
QY 1174 -----ACTCTAGAGCTGCAACCCCAACCATCGCTCCCGCGCACAAA----- 1215
Db :||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
485 ArgHisGlyLeuLeuLeuArgAlaProValProGlyGlnProHisGlnProArg 504
QY 1216 -----AGCAAAAGGCTGACCGGACGCTCCGTAACACGCGCTGCGCAAAAC 1263
Db :||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
505 AlaGluArgGlyArgAlaGlnProGlyArgGluLeuProProGlyProArgLeuGlyGln 524
QY 1264 CTTTCATTGGACGGCAA-----ACGCATCTTCGCGCTCATTCAA 1305
Db :||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
525 AspArgArgGlyAspArgHisProGluAlaHisValValHisLeuHisArgGlyAlaVal 544
QY 1306 CCT----- 1308
Db :||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
545 ProGlyArgGlyProAlaProProArgGlyGluHisGlnGlyValGlyGluGluHisArg 564
```

```
QY 1309 -----CGACCC 1314
Db 565 AspProGlyGlyGlnGluGlyAlaAspHisGluProLeuGlyArgAlaLeuGlnArgPro 584
QY 1315 -----CGTTCCGACCGCTCAATACCGAAAT 1341
Db 585 LeuGlnArgGluGlyAlaAspGlnArgHisArgProArgGlyArgValHisValArgGly 604
QY 1342 GGGCGTCTCATCGAAAGCCC-----CAAAATCGCAGAACACAGATCGAGCG 1386
Db :||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
605 GlyArgGlyGlnArgGlnProAlaAlaAspAlaGluAlaArgArgAlaGlyGlyPro 624
QY 1387 CACCTCTCCCATACACACCGCAATACGCTACGCGGT----- 1425
Db :||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
625 ArgProGlnGlnArgArgAlaArgSerGlySerProProCysSerProArgSerProGly 644
QY 1426 -----TACCTCGCAACAAACACACCGCTCGCAATGCGACA 1461
Db 645 SerArgArgSerSerAlaArgCysCysProArgArgTrpArgProAlaTrpArgArg 664
QY 1462 TCCCGCCAC---CCGAAAAACCTACCCGAAACGA 1491
Db 665 ArgGlyHisArgProArgGlyGluProAspArg 675

RESULT 29
US-11-096-568A-23618
; Sequence 23618, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nikolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 23618
; LENGTH: 375
; TYPE: PRT
; ORGANISM: Zea mays subsp. mays
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(375)
; OTHER INFORMATION: Ceres Seq. ID no. 12414168
US-11-096-568A-23618

Alignment Scores:
Pred. No.: 3e-08 Length: 375
Score: 235.50 Matches: 119
Percent Similarity: 35.2% Conservative: 46
Best Local Similarity: 25.4% Mismatches: 149
Query Match: 8.3% Indels: 156
DB: 11 Gaps: 20

US-10-665-990A-13 (1-1561) x US-11-096-568A-23618 (1-375)
QY 157 CTGTGACAAACATCTCGCAATCCGSCACACCCCTCATAAACACGGGCTATCCGACATCTA 216
Db 4 ProAlaArgArgGlnHisSerProThrHis-----GlyValArgSerArg 18
QY 217 CTGTCTCAGACCCCGCCGAGGCGCTTGCAGCGCGCGCGCGCTTATCGATCTGCCGA 276
Db 19 AlaAlaGlyArg-----ArgGlyArgValProGlyArgAlaValArg 32
QY 277 ACACAGCCTCATTTTGCAATACTACATTTGGCGCAACGACATTTCCGCGAGGCTGCTGT 336
Db 33 ValGlnArgArg----- 36
QY 337 CAACCTCATGTACTCTTCCGCGAGAACGCGCGGTGGCGGTAGCCCTGCTGTTGGACGACAA 396
Db 37 -----HisGlyGlyAlaAlaArgAlaAlaGlyValArgAspAlaGlyAlaGlyArg 54
QY 397 CAACA-----CGCGCGGGTTGG---ACGATCTCC 422
```

```
Db 55 HisGlyArgLeuGluGlyArgAspValArgGlyProAlaGlyTipArgSerAlaSer 74
QY 423 TGCTCGCCTCGACAGCATCCCAATATCGAAGTCGGCTGTTCACCCCTTCGTCCTAC 482
Db 75 ThrSerProGlyThrAlaGlyProArgProGlyGlyAlaCysArgSer---SerSerThr 93
QY 483 GCAATGGCGCGCACTCGCTACCTGACCGACTTCCCGCGCTCAACCGCGCATGCACA 542
Db 94 ThrThrAlaAlaSerAlaSerGlyPro---AlaProGlyProThrAlaArgThrThr 112
QY 543 ACAATCTTTACCGCGGACAAACCGGCACCATATCGCGGACGCAATATCGCGCAGC 602
Db 113 AlaCysAlaTtpProProThrTtpAlaArgSerTtpArg----- 126
QY 603 AATACTTCAAAGTCGGTGAGGACACCGTTTCGCGGACCTGGACATCTCGCCACCGCA 662
Db 127 -----ArgThrThrAlaSerArgProSerThrAlaSerProProPro 140
QY 663 CGCTCGTCGGGAAATGATCGACGACCTTCGACCGCTACTGGGCAAGCCATTCCGCCACA 722
Db 141 SerThrThrAlaArgArgSerCysGlyTtpArgGlyArgAlaAlaGlyThrProGly 160
QY 723 ACGCACCGCATCATCGCGAGCGGCAACATCGGCAAGGCTTCAAGCACTCGGATACA 782
Db 161 SerProArgProProThrSerAlaAlaSerSerCysArgAla----- 174
QY 783 ACGAGAAACATCCAGACACCGCTCTCGCTACCGGAAACCGTCGAACAGTCGCCCC 842
Db 174 ----- 174
QY 843 TCTACAAAAAATACAGACGGGACGATCGACTGCGAGCGGTCCAAACCGCGCTGATCA 902
Db 175 -----ThrArgProAlaAlaProSerArgThrThrSerProCys----- 187
QY 903 GCGACACCTTGCAAAAGGACTCGACCGGACCGCGCAACCGCCGATG-----CCG 956
Db 188 -----AlaSerValAlaArgProArgThrTtpProPro 198
QY 957 GGAGGCTGCAAGCGCTCAACAGCCCGAAAGAGGTCTATCTGTTTCAACCTATT 1016
Db 199 SerProCysGlyAspThrSerSerSerCysHisSerSerAlaAlaTtp----- 214
QY 1017 TCGTCCCTTACAAAATCCGGCACAGACGCACTGGCAAACTGGTCAGGACGGCATAGC 1076
Db 214 ----- 214
QY 1077 TTACCGTCTGACCAACTCGCTACAGGCGACCGAGTTGCGCGCTCCATTCCGGCTACG 1136
Db 215 -----ArgAlaArgGlyArgArgProSerAlaProProThrLeuSerSerThr 230
QY 1137 TCAATACCGAAACCGCTGCTCAAGCGGCATCAACTCTACGAGTGC----- 1187
Db 231 GlyPro-----SerThrThrAlaThrGlyAlaCysArgCysArgAla 245
QY 1188 ---AACCCACCATCGCTCCCGCCACAAAAGACAAAGGCTGACCGGCA----- 1235
Db 246 ProArgProThrThrProTtpPro-----ThrArgSerAlaProAlaArgArgArg 262
QY 1236 ---GCTCGGTAAACGAGCTGATCGCAAAACCTTCATTGTGGAGCGGCAACGCACTTCA 1292
Db 263 TtpThrProSerSerArgProProTtpSerSerSerAlaAlaAlaThrSerSerThr 282
QY 1293 TCGGCTCATTAACCTGACCCCC-----GTTCCGACGCTCAATACCGAAATGGCG 1346
Db 283 ThrAlaProSerThrThrProGlySerGlyProGlySer---ProSerTtpSer 301
QY 1347 TCGTCATGAAAGCCCAAAATCCAGAAAGATGGAGCGCACCTCGCGATACCAACAC 1406
Db 302 GlyThrSerThrAlaSerSerThrAlaSerSerProSerThrProGlyProProPro 321
QY 1407 CCG-----AATACGCTACCGCGTTA---CCCTCGACA 1436
|||
:::|||||
```

```
Db 322 ProSerSerCysAlaSerSerSerAlaSerSerThrProThrAlaAlaSerThrSerThr 341
QY 1437 AACACACCGCGCTGCAATGGCAGCATC 1463
Db 342 AspThrSerValSerAsnSerSerIle 350
RESULT 30
US-10-517-696-111
; Sequence 111, Application US/10517696
; Publication No. US20060051759A1
; GENERAL INFORMATION:
; APPLICANT: diaDexus, Inc.
; APPLICANT: Salceda, Susana
; APPLICANT: Macina, Roberto A.
; APPLICANT: Turner, Leah R.
; APPLICANT: Sun, Yongming
; APPLICANT: Liu, Chenghua
; TITLE OF INVENTION: Compositions and Methods Relating to Breast Specific Genes and
; FILE REFERENCE: DEX-0432
; CURRENT APPLICATION NUMBER: US/10517,696
; CURRENT FILING DATE: 2004-12-13
; PRIOR APPLICATION NUMBER: US 60/389,327
; PRIOR FILING DATE: 2002-06-14
; NUMBER OF SEQ ID NOS: 171
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 111
; LENGTH: 1256
; TYPE: PRT
; ORGANISM: Homo sapien
US-10-517-696-111
Alignment Scores:
Pred. No.: 3,81e-08 Length: 1256
Score: 235.00 Matches: 146
Percent Similarity: 36.0% Conservative: 55
Best Local Similarity: 26.2% Mismatches: 238
Query Match: 8.2% Indels: 119
DB: 24
Gaps: 24
US-10-665-990A-13 (1-1561) x US-10-517-696-111 (1-1256)
```

```
QY 93 CATGGTTGCCCGCCACTCGAAGACGAGCGGAAAGCGCTCATTTCAATATCTTCCAAACCTG 152
Db 156 HisGlyValThrSerAlaProAspThrArgProAlaProGlySerThrAlaProProAla 175
QY 153 TCCTCTGGACACATCTCGCAATCCGGCACCCCTCATACAAACGGGTATCCGACA 212
Db 176 HisGlyValThrSerAlaProAspThrArgProProGlySerThrAlaProProAla 195
QY 213 TCTACCTGTCGACGACCCCGCC---ACGAAGCCC-----TTGCCGCGCGCG 254
Db 196 HisGlyValThrSerAlaProAspThrArgProAlaProGlySerThrAlaProProAla 215
QY 255 CCGCGCTTATCGAATCTGCGCAACACACAGCTCGATTGCAATCTACATTTGGCGCAACG 314
Db 216 HisGlyValThrSerAlaProAspThr-----ArgProAlaProGlySerThr 231
QY 315 ACATTTCCGCA---GGCTGCTGTTCACCTCATGTACCTTGGCG-----CAGAAC 362
Db 232 AlaProProAlaHisGlyValThrSerAlaProAspThrArgProAlaProGlySerThr 251
QY 363 GCGCGCTGCGGTACGCTGCTGTCGACGACACACACGCGCGGTGGACCATCTCC 422
Db 252 AlaProProAlaHisGlyValThrSerAlaProAspThrArgProAlaProGlySerThr 271
QY 423 TGCTCGCGCTCG-----ACAGCCATCCCAATATCGAAGTGGCGCTGTTCACCCCT 473
Db 272 AlaProProAlaHisGlyValThrSerAlaProAspThrArgProAlaProGlySerThr 291
QY 474 TCGTCTTACCAATGGCGCGCATCGCT-----ACC 506
Db 292 AlaProProAlaHisGlyValThrSerAlaProAspThrArgProAlaProGlySerThr 311
|||
:::|||||
```



```
Db 240 ProAlaLeuLeuHisArgProHisHisGlyProPro----- 252
QY 556 CGCCGCAACCGCCACCATACTCGCGGACGCAATATCGCGACGATACTTCAAAGT 615
Db 253 GlnArgAlaGlyArgHisArgArgGlyGlyArgArgGlyValGlnAsp 272
QY 616 CGGTGAGGACACCGTTTTCGC-----CGACCT---GGACATCTCTCGCCACCGG 660
Db 273 Arg---ArgHisArgGlyArgLeuLeuGlnAlaGlnProGlnGlyProArgHisArg 291
QY 661 CAGCCTCGT-----CGCGAAGTATCGCAGCTTCGACCG 696
Db 292 GlnArgHisValArgGlyLeuArgAlaArgGlyHisArgAspValArgGlnArgPro 311
QY 697 -----CTACTCGGCAAGCCATTCGCG----- 717
Db 312 GlyArgProValGlyGlyProValArgArgLeuLeuArgGlyHisGluArgGlnAlaArg 331
QY 718 -----CCACAA----- 723
Db 332 ValHisGlyProProAspProGlnAlaGluAlaProProGlyValHisArgGlyArgGly 351
QY 724 CGCCACGCGCATCATCG-----CAGCGGCAACATCGCAAGGTCTTCAAGCACTCGG 777
Db 352 HisHisGlyAlaHisProGlyTrpGlnLeuLeuHisGluAlaGlyGlnGluArg 371
QY 778 -----ATACAACGACGAAACATCCAGACACGCGCTCTCGCTACCGCGAAAC 825
Db 372 AlaGlyProAlaAlaGluAlaGlnAlaGlyGlnValArgAlaPro----- 386
QY 826 CGTCGAACAGTCGCGCCCTCTACCAAAAAATACAGACGGGCGCATCGACTGCGAGCGT 885
Db 387 -----HisValAlaAlaVal 391
QY 886 CCAACACCGCTGATCAGGACACCCC----- 912
Db 392 AlaGlyProProAspArgGlyHisProArgArgHisGlnValHisArgAlaArgGlyGln 411
QY 913 TGCAAAAGACTTCGACCG-----CGACCGCGCAAAACCGCGATTGCGGGAGGCT 963
Db 412 LeuArgGluArgGlnProGlyHisArgArgProProArg----- 424
QY 964 GCAAGACGGCTCAACAGCCGCGAAAGGCTGTATCTGTGTTTCAACCCTATTTCTGCTCC 1023
Db 425 -----GlnGlyAlaAlaArgArgGlnLeuProGlyHisPro----- 436
QY 1024 TACAAATCCGCGCAGCGACTCGGCAAACTGTGTCAGGACGGCATAGAGTTACCGT 1083
Db 437 -----HisArgValHisGlyGlnArgProProArgHisArgGlnHisArg 452
QY 1084 CTGTACCAACTCGTACAGGCGACCGACGTTGCGCGCTCCATTCCGCGTACGTCAAATA 1143
Db 453 GlnAlaHisValArgAlaValLeuArgAlaArgGlnArgValLeuGlnArgAlaHis 472
QY 1144 CGGAAACCGTGTCTAAAGCGCGCATCAA----- 1173
Db 473 LeuGlnProGlyArgGlnProGlnProGlnProGlyLeuArgLeuGlnGlyHisArgAsp 492
QY 1174 -----ACTCTAGAGCTGCAACCCACCATCGCTCCGCGCACAAAGACAA 1221
Db 493 ArgHisGlyLeuLeuArgAlaProValProGlyGlnProHisHisGlnProArg 512
QY 1222 -----AGCCTGACCGGACGTCGTAACCGCTGCATGCCAAAC 1263
Db 513 AlaGluArgGlyAlaAlaGlnProGlyArgGluLeuProGlyProArgLeuGlyGlnGlu 532
QY 1264 CTTCACTTGGACGGCAA-----ACGCACTTTCATCGCTCATT--- 1302
Db 533 AspArgArgGlyAspArgHisProGluAlaHisValValHisLeuHisArgGlyAlaVal 552
QY 1303 -----CAACCTCGACCCCGTTCCGACGGCTCAATACCGAATGGCGTCGTCAT 1353
Db 553 ProGlyArgGlyProAlaProPro---ArgGlyGluHisGlnGlyValGlyGluGluHis 571
```

```
QY 1354 CGAAGAGCCC----- 1362
Db 572 ArgAspProGlyGlyGlnGlnGlyAlaAspHisGluProLeuGlyArgAlaLeuGlnArg 591
QY 1363 -----CAAAATCGCAGAACAGATGGAGCGCACCTCGCCGATACCCACCCGAATACGC 1416
Db 592 ProLeuGlnArgGlyGlyAlaAspGlnArgHisArgProArgGlyArgValHisValArg 611
QY 1417 CTACCGGTTACCTTCGACAAACACACCGCTCGAATGGCAGCATCCCGCCACCCGAAA 1476
Db 612 GlyGlyArgGlyGlnArgGlnProAlaAlaAspAlaGluAlaAlaArgAlaGlyGly 631
QY 1477 AACCTACCCGAAACCCGAAAGC-----CAAACCTTTGGAACCGCATCGC 1521
Db 632 ProArgProGlnGlnArgArgArgGlyAlaGlyAlaLeuHisValLeuGlnAspHisGln 651
QY 1522 CGCAAAATCTATCCCT 1539
Db 652 ValArgGlyGlyAlaPro 657

RESULT 32
US-11-096-568A-28958
; Sequence 28958, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nikolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptide.
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 28958
; LENGTH: 1480
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(1480)
; OTHER INFORMATION: Ceres Seq. ID no. 3095666
US-11-096-568A-28958
```

Alignment Scores:

Pred. No.:	5,7e-08	Length:	1480
Score:	232.50	Matches:	149
Percent Similarity:	32.5%	Conservative:	43
Best Local Similarity:	25.3%	Mismatches:	219
Query Match:	8.2%	Indels:	180
DB:	11	Gaps:	29

US-10-665-990A-13 (1-1561) x US-11-096-568A-28958 (1-1480)

```
QY 34 CCGCATGAAACACGACGCTCATTTCCCTTTTATGCTCTCTCTCTCTCTCTCTCTTC 93
Db 95 ProLysProThrLysProHisProHisProLysProThrLysPro----- 112
QY 94 ATGTTGCCCTGGAAGAACGGACGGAAGCCGTCATTTCATCTTCCAAACCTGT 153
Db 113 ProProHisProLysProArgProHisProLysPro-----ProAsnVal 127
QY 154 CTTCTGGCAACATCTCGCAATTCGGCA-----CACCCCTCATAAACACGGCTATC 207
Db 128 LysProHisProHisProLysProProThrLysProHisProHisProLysProProThr 147
QY 208 CGACATCTACTGCTCGACGACCCCCACG----- 236
Db 148 LysHisHisProHisProLysProProThrLysProProLysProProSerVal 167
QY 237 AAGCCCTTGCG-----CCCGCGCGCTTATCGAATCTGCGAACACAGCCCTCGATT 290
Db 168 LysProProSerThrProLysProProThrThrAsnProPro----- 183
```

```
QY 291 TGAATACTACATTTGGCGCAGACATTTTCGGCAGGCTGCTTCAACCTCATGTACC 350
Db ::::|
184 ---SerThrProGlnProProThrHisLysProProProCysThrProThr----- 199
QY 351 TTGGCGCAGAACGGCGTGCGGTAGCGCTGCTGTGGAGCACAACAACACGCGGGGT 410
Db ::::|
200 ---ProProValAlaSerProProMetAla-----ThrProThrGln----- 213
QY 411 TGGACGATCTCTGCTCGCCCTCGACAGCCATCCCAATATCGAAGTGGCGCTGTTC AAC 470
Db ::::|
214 ---MetProProIleAlaThrProProIleAlaLysSerProValAlaThr 229
QY 471 CTTCTGCTCTAGCAAAATGGCGGCACCTCGGTACTGACCGACTCC----- 518
Db ::::|
230 ProPro-----IleAlaThrProProThrAlaThrProProIle 242
QY 519 -----CCGCTCAACCGCGCATGCACAAATCCT----- 551
Db ::::|
243 ThrIleProProValAlaThrProProIleThrThrProProIleAlaAsnProProIle 262
QY 552 ---TTACCGCGACAAACCGCCACCATCTCGCGGACGCAATATCGCGCAGCAATACT 608
Db ::::|
263 IleMetProProIleAlaThrProProValAlaAlaProProIleThrAsnProProIle 282
QY 609 TCAAG---TCGTTGAGGACACGCTTTCCGCGACCTGGACA----- 647
Db ::::|
283 SerLysProProValThrThrPro-----ProThrThrThrProProIleAlaLysPro 300
QY 648 -----TCCTCGCCACCGGAGCGTCTCGCGGAAGTATCGCAGC 686
Db ::::|
301 ProIleAlaThrProProIleSerThrProProAlaAlaThrProProAlaAlaThrPro 320
QY 687 ACTTCGACCGTACTGGCAAGCC-----ATTCG 716
Db ::::|
321 ProIleThrThrLeuProProAlaLysProProValAlaIleSerProIleValThrPro 340
QY 717 CCCACACCGCGCGCATCTCCGAGCGGAACATCGCGAAGGTCTTCAAGCACTCG 776
Db ::::|
341 ProValThrProIleAlaGlnProProValAlaThr----- 352
QY 777 GATACAAACGAGAAACATCCAGACGCGCTCTCGCTACCGCGAAACCGTCGACAGT 836
Db ::::|
353 -----ProProThrAlaThrProProValAlaThrProProIleAla 366
QY 837 CGCCCTCTACAAAATAACAGCGGACGATCGACTGGCAGAGCGTCCAAACCGCGC 896
Db ::::|
367 ThrProProThrSerLys-----SerProIleSerThrProPro 379
QY 897 TGATCAGCAGACCCCTGCAAAAGGACTCGACGCGACCGCGCAACCGCGCATGCG 956
Db ::::|
380 IleSerGluSerProValAlaThrProProThrAlaThrSerProIleLysThrProPro 399
QY 957 GGAGGTGCAAGAGCGGTCAAAACAGCGCAAAAGCGTCTATCTGGTTTCAACCTATT 1016
Db ::::|
400 -----ProAlaLysProProValAlaThrProProIle 410
QY 1017 TCGTCCCTCAAAATCCGCGACAGCGCACTGGCAAACTGTGCGAGGACGGATAGAC 1076
Db ::::|
411 Ala-----LysSerProIleAlaThr----- 417
QY 1077 TTACCGTCTGACCAACTCGCTACAGGCGACGAGTTCGCCCGCTCCATTCGGGTACG 1136
Db ::::|
418 -----ProProThrAlaThrProProVal---AlaThr 427
QY 1137 TCAAAATACGAAACCGC-----TGCTCAAGCGCGGCATCAAACTCTACGAGCTGCAAC 1190
Db ::::|
428 ProProIleGluLysProProValAlaThrProProThrThrThrProProThrAlaThr 447
QY 1191 CCAACCATCGCTCCCGCGCACAAGACAAAGGCGCTGACCGCAGCTCCGTAAACAGCC 1250
Db ::::|
448 ProProValAlaLysProProValGluThrPro-----ProIleAlaThrProProThr 465
QY 1251 TGCATGCCA-----AAACCTTCATTGTGGAGCGGCAACGCA 1286
```

```
Db 466 AlalysProProIleSerThrProProIleSerLysProProValAlaThrProProIle 485
QY 1287 TCTTCATCGGTCTATTC AACCTCGACCCCGTT-----CCGCACGGC 1328
Db ::::|
486 AlaThrProProIleThrThrProThrProValLysProProValAlaThrProProLeu 505
QY 1329 TCATACCGAAATGGCGTCG-----TCATCGAAAGCCCAAAATCG 1370
Db ::::|
506 AlaIleProProValAlaLysProProValValThrProProThrAlaThrProProIle 525
QY 1371 CAGAACAGATGGAGCGACCTCGCGCATACACACCGCAATACCGCTACCGCTTACCC 1430
Db ::::|
526 AlaThrProProIleAlaLysSerProValAlaThrPro-----ProThrAlaThrPro 543
QY 1431 TCGACA-----AACACACCGCC 1448
Db ::::|
544 ProValAlaThrProProIleAlaLysProProValValThrProProThrThrThrPro 563
QY 1449 TGCATGGCAGATCCCGCCACCGCAAAACCTACCCGAAACGACCGACCGCAACTTT 1508
Db ::::|
564 ProThrAlaThrProProValAlaLysProProValAlaThrProProIleAlaThrPro 583
QY 1509 GAAACGATCGCGCAAAATCCTATCCC 1538
Db ::::|
584 ProThrAlaLysProProIleSerThrPro 593

RESULT 33
US-10-793-626-1000
; Sequence 1000, Application US/10793626
; Publication No. US20050255478A1
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PUS480US
; CURRENT APPLICATION NUMBER: US/10/793,626
; PRIORITY FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1000
; LENGTH: 442
; TYPE: PR1
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: amino acid sequence
; FEATURE:
; NAME/KEY: MOD RES
; LOCATION: (442)
; OTHER INFORMATION: variable amino acid
US-10-793-626-1000

Alignment Scores:
Pred. No.: 5,248-08 Length: 442
Score: 232.00 Matches: 89
Percent Similarity: 37.3% Conservative: 57
Best Local Similarity: 22.8% Mismatches: 157
Query Match: 8.1% Indels: 88
DB: 9 Gaps: 9

US-10-665-990A-13 (1-1561) x US-10-793-626-1000 (1-442)
QY 212 ATCTACCTGCTCGACAGCCCCACGAGCCCTTGGCCCGCGCCCTTATCGAATCT 271
Db ::::|
129 IleAspLeuPheThrAspGlyHisLysLeuTyrGluLysValLeuGluAspIleTyrAsn 148
QY 272 GCGCAACACAGCCTCGATTGCAATACATACATTTGGCGCAACGACATTTCCGCGAGCTG 331
Db ::::|
149 AlaGlnAspTyrIleHisLeuGluTyrTyrThrPheGluLeuAspGlyLeuGlyLysArg 168
QY 332 CTGTTTCAACCTCATGTACTTGTCCGAGAACCGCGCGTGTACGCTGTGTGGAC 391
```



```
Db 169 IleLeuAspAlaLeuGluThrLysLeuLysGluGlyLeuGluValLysLeuLysLeuTyrAsp 188
QY 392 GACAAACACACGCGGGTGGACGATCTCTCTGCTCGACGACGATCCCAATATC 451
Db 189 AspValGlySerLys-----
QY 452 GAAGTCGCGCTGTTCAACCCCTTCGTCCTACGAAATGGCGCGACTCGGCTACCTGACC 511
Db 194 LysValArgLeuSerLys-----PheLysHisPheArgAlaLeuGlyGlyVal 210
QY 512 GAC-----TTCCCC-----CGCCTCAACGCGCGCATGCAC 541
Db 211 GluAlaPheProSerLysValProLeuIleAsnPheArgMetAsnAsnArgAsnHis 230
QY 542 AACAAATCTTTACCGCGCAACACGCGCCACCATCTCGCGGAGCGCAATATCGCGAC 601
Db 231 ArgLysIleIleIleAspGlyGlnIleGlyTyrValGlyGlyPheAsnValGlyAsp 250
QY 602 GAATATCTTCAAGTCGGTGAGACACCGGTTTTCGCGGACCTGGACATCTCGCCACCGGC 661
Db 251 AspTyrLeuGlyLysLeuGlyTyrTrpArgAspThrHisThrArgValGlnGly 270
QY 662 AGCGTCGTCGCGCAAGTATCGACGACTTCGACCGCTACTGGGCAAGCATTCGCGCCAC 721
Db 271 GluCysIleAspAlaLeuGlnLeuArgPheIleLeuAspTrpAsnSerGlnSer----- 288
QY 722 AACGCGACGCGCATCTCGCGAGCGCAACATCGCGCAAGGGTCTTCAAGCATCTCGATAC 781
Db 288 -----
QY 782 AACGACGAAACATCAGACACGCGCTCTCGCGTACTCGCGCAACACGCGTCAACGATCGCGCC 841
Db 289 -----HisArgProGlnPheLysPheAspGln 297
QY 842 CTCATACCAAAATACACAGCGGACGATCGACTGGGAGAGCGTCCAAACCGCGCTGATC 901
Db 298 LysTyrPheProLysLysAsnGly-----AspLysGlyAsnAlaIleGlnIleAla 315
QY 902 AGCGACACCCCTGCAAAAGGACTCGACCGCGACCGCGCCGCAACCGCGCATTCGCGGAGG 961
Db 316 SerSerGlyProAlaPheAspLeuHisGln-----IleGluTyrGly 329
QY 962 CTGCAAGACGCGCTCAAAACAGCCGCTATCTGTTTCAACCTATTTCTGTC 1021
Db 330 TyrThrLysMetIleMetSerAlaLysLysSerIleTyrLeuGlnSerProTyrPheIle 349
QY 1022 CTTACAAATCTGGGACACGACGACTGGCAAACTGGGTGCGAGGCGCATAGACGTTACC 1081
Db 350 ProAspGlnSerTyrIleAsnAlaLeuLysMetAlaAlaAsnSerGlyValGluValAsn 369
QY 1082 GTCCTTGACCAACTCGCTACAGCGGACCGACGCTGCGCGCGCTCCATTCGCGGTACGTCAA 1141
Db 370 LeuMetIleProCys-----LysProAspHisProPheValTyrTrpAlaThrPheSer 387
QY 1142 TACGAAACCGCTGCTCAAAAGCGGACATCAAACTCTACGAGCTGCAACCAACCATGCC 1201
Db 388 AsnAlaAlaAspLeuLeuAspSerGlyValAsnIleTyrThrTyrGlnAsnGlyPhe---- 406
QY 1202 GTCCCCCGCACAAAGAGCGCTGACCGCGAGCTCGTAACCGACGCTCGATCGCATGCCAAA 1261
Db 407 -----IleHisSerLys 410
QY 1262 ACCTTCATTGTGGACGCAACGATCTTTCATCGGCTCATTCACCTCGACCCCGGTTC 1321
Db 411 IleLeuMetIleAspAspGluIleSerSerIleGlySerAlaAsnMetAspPheArgSer 430
QY 1322 GCAGCGCTCAATACGAAATGGCGTCTGTCATC 1354
Db 431 PheGluLeuAsnPheGluValAsnAlaPheIle 441
```

RESULT 34

US-11-050-857-487

```
/ Sequence 487, Application US/11050857
/ Publication No. US20060040278A1
/ GENERAL INFORMATION:
/ APPLICANT: Compugen Ltd
/ TITLE OF INVENTION: NOVEL NUCLEOTIDE AND AMINO ACID SEQUENCES, AND ASSAYS AND METHODS
/ FILE REFERENCE: 1847.1005
/ CURRENT APPLICATION NUMBER: US/11/050.857
/ CURRENT FILING DATE: 2005-01-27
/ NUMBER OF SEQ ID NOS: 1150
/ SEQ ID NO 487
/ LENGTH: 1255
/ TYPE: PRF
/ ORGANISM: Homo sapiens
US-11-050-857-487
```

```
Alignment Scores:
Pred. No.: 7,01e-08 Length: 1255
Score: 231.00 Matches: 145
Percent Similarity: 35.8% Conservative: 55
Best Local Similarity: 26.0% Mismatches: 239
Query Match: 8.1% Indels: 119
DB: 11 Gaps: 24
```

US-10-665-990A-13 (1-1561) x US-11-050-857-487 (1-1255)

```
QY 93 CATGTTGCCCCCACTGGAGACGAGCGGAAAGCGTCATTTCAATACTTTCCAAACCTG 152
Db 436 HisGlyValThrSerAlaProAspThrArgProAlaProGlySerThrAlaProProAla 455
QY 153 TCCTCTCGGACAACTCTGCAATCCGCGACACCCCTCATAAACAAAGGGTATCCGACA 212
Db 456 HisGlyValThrSerAlaProAspThrArgProAlaProGlySerThrAlaProProAla 475
QY 213 TCTACCTGTCGACGACCCCC---ACGAAGCCC-----TTGCGCGCCGCG 254
Db 476 HisGlyValThrSerAlaProAspThrArgProAlaProGlySerThrAlaProProAla 495
QY 255 CCGCCCTTATCGAATCTGCGACACACACGCTCGATTGCAATATCTACATTTGGCGCAACG 314
Db 496 HisGlyValThrSerAlaProAspThr-----ArgProAlaProGlySerThr 511
QY 315 ACATTTCCGGCA---GGCTGCTGTTCACACCTCATGTACTTTCGCG---CAGAAC 362
Db 512 AlaProProAlaHisGlyValThrSerAlaProAspThrArgProAlaProGlySerThr 531
QY 363 GCGCGCTGCGGTACGCTGCTGTGGACGACAAACACGCGCGGTGGACGATCTCC 422
Db 532 AlaProProAlaHisGlyValThrSerAlaProAspThrArgProAlaProGlySerThr 551
QY 423 TGCTCGCCCTCG-----ACAGCCATCCCAATATCGAAGTCGCGCTGTTCAACCCCT 473
Db 552 AlaProProAlaHisGlyValThrSerAlaProAspThrArgProAlaProGlySerThr 571
QY 474 TCGTCTTACCAATATGCGCGCACTCGGCT-----ACC 506
Db 572 AlaProProAlaHisGlyValThrSerAlaProAspThrArgProAlaProGlySerThr 591
QY 507 TGACCGACTTCCCGCGCTCAACGCGCGATGCACAAACAATCTTTTCGCGCGCAAC 566
Db 592 AlaProProAlaHisGlyValThrSerAlaProAspThrArgProAlaProGlySerThr 611
QY 567 GCGCCACCATACTCGCGGAGCAATATCGCGACGAATCTTCAAGTCGTCGAGGACA 626
Db 612 AlaProPro-----AlaHisGlyValThrSerAlaProAspThrArgProAla 627
QY 627 CCGTTTTTCGCGGACCTCGGACATCTCGCCACCGCGCGCTCGTCGCGGAGATATCGCAGC 686
Db 628 ProGlySer-----ThrAlaProProAlaHisGlyValThrSerAlaProAsp 643
QY 687 ACTTCGACCGCTACTGGGCAAGCCATTTCGCCCC----- 719
Db 644 ThrArgProAlaProGlySerThrAlaProProAlaHisGlyValThrSerAlaProAsp 663
```

```
QY 720 ACACGCCGCGCATCTCCGACGGCGAACAATCGGCAAGGGTCTTCAAGCACTCGGAT 779
Db 664 ThrArgProAlaProGlySerThrAlaProAlaHisGlyValThrSerAlaProAsp 683
QY 780 ACA-----ACGACGAAACATCCAGAC 800
Db 684 ThrArgProAlaProGlySerThrAlaProProAlaHisGlyValThrSerAlaProAsp 703
QY 801 ACGCGCTCTGCG-----GCTACCGGAAACCGTCCGAAC-----AGTCGCCCTCT 845
Db 704 ThrArgProAlaProGlySerThrAlaProProAlaHisGlyValThrSerAlaProAsp 723
QY 846 ACACAAAAATACAGACGGACGATCGATCGGACAGCGTCCAAACCGCCTGATCAGCG 905
Db 724 ThrArg-----ProAlaProGlySerThrAlaProProAlaHisGlyVal 738
QY 906 ACACCCCTGCAAAAGGACTCGACCGGACCGCGCAACCGCGGATTCGGGAGGCTGC 965
Db 739 Thr-SerAlaProAspThrArgProAlaProGlySerThrAlaProProAlaHisGlyVa 758
QY 966 AAGACGCGTCAACACAGCCGCAAAAGCGTCTATCTGCTTTTCACCCCTATTTTCGTCCTA 1025
Db 758 lThrSerAlaProAspThrArgPro----- 766
QY 1026 CAAATCCGCGCAGACGCACTGGCAAACTGGTCGACGCGCATAGACGTACGTC 1085
Db 767 -----AlaProGlySerThrAlaProProAlaHis-GlyValThr-Ser 780
QY 1086 TGACCAACTCGTACAGGCGACG-----AGTTGCGCGCTGCATTCGCGCTACGTCA 1139
Db 781 AlaProAspThrArgProAlaProGlySerThrAlaProProAlaHisGlyValThrSer 800
QY 1140 AATACCGAAACCGC-----TGCTCAAGCCGCGCATCAAACTCTACGAGC 1184
Db 801 AlaProAspThrArgProAlaProGlySerThrAlaProProAlaHisGlyValThrSer 820
QY 1185 TGAACCCAAACGTCGCTCCCGCCCAAAAGACAAAGCGCTGACGGCA----- 1235
Db 821 AlaProAspThrArgProAlaPro-----GlySerThrAlaProProAlaHisGlyVal 838
QY 1236 -----GCTCGTAACGACGCTGCATGCCAAACCTTCATG 1271
Db 839 ThrSerAlaProAspThrArgProAlaProGlySerThrAlaProProAlaHisGlyVal 858
QY 1272 TGAACGCGCAACGATCTTCATCGGCTCATCAACCTCGACCCCGCTCCGACGCG--- 1328
Db 859 ThrSerAlaProAspThrArgProAlaProGlySerThrAlaProProAlaHisGlyVal 878
QY 1329 ----TCAATACG-----AAATGGGCGTGTCTATCGAAAGCCCAAAATCGCAGACAGA 1379
Db 879 ThrSerAlaProAspThrArgProAlaProGlySerThrAlaProProAlaHisGlyVal 898
QY 1380 TGAAGCGCACCTCGCGGATACACACCGCGATACGCTACCGCGTTACCGTCGACAAAC 1439
Db 899 ThrSerAlaProAspThrArgProAlaProGlySer---ThrAlaProProAlaHisGly 917
QY 1440 ACACCCGCTGCAATGGCAGCATCCCG-----CCACCCGAAAAACCT 1481
Db 918 ValThrSerAlaProAspThrArgProAlaProGlySerThrAlaProProAlaHisGly 937
QY 1482 ACCCGAACGACCG-----ANGCCAACTTTGGAAACGCAATCGCGC 1523
Db 938 ValThrSerAlaProAspAsnArgProAlaLeuGlySerThrAlaPro 953

RESULT 35
US-11-043-806-398
; Sequence 398, Application US/11043806
; Publication No. US20060051774A1
; GENERAL INFORMATION:
; APPLICANT: Comugen Ltd
; TITLE OF INVENTION: Novel Nucleotide and Amino Acid Sequences, and Assays and Methods
; THEREOF FOR DIAGNOSIS OF PROSTATE CANCER
```

```
; FILE REFERENCE: 1847.1003
; CURRENT APPLICATION NUMBER: US/11/043.806
; CURRENT FILING DATE: 2005-01-27
; NUMBER OF SEQ ID NOS: 575
; SEQ ID NO 398
; LENGTH: 1255
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-043-806-398
```

```
Alignment Scores:
Pred. No.: 7,01e-08 Length: 1255
Score: 231.00 Matches: 145
Percent Similarity: 35.8% Conservative: 55
Best Local Similarity: 26.0% Mismatches: 239
Query Match: 8.1% Indels: 119
DB: 11 Gaps: 24
US-10-665-990A-13 (1-1561) x US-11-043-806-398 (1-1255)
```

```
QY 93 CATGTTCCCGCCCTGGAAGACGCGAAGCGCTCATTTCAATACTTCCAAACCTG 152
Db 436 HisGlyValThrSerAlaProAspThrArgProAlaProGlySerThrAlaProAla 455
QY 153 TCCTCTCGACACATCTGCAATCCGGCACACCCCTCATAAACACGGCTATCCGACA 212
Db 456 HisGlyValThrSerAlaProAspThrArgProAlaProGlySerThrAlaProAla 475
QY 213 TCTACTCTCGACACACCCCGC---ACGAAGCGCC-----TTGCGCGCCGCG 254
Db 476 HisGlyValThrSerAlaProAspThrArgProAlaProGlySerThrAlaProAla 495
QY 255 CGCGCTTATCGAATCTCGGAAACACACGCTCGATTGCAATACTACATTTGGCGCAACG 314
Db 496 HisGlyValThrSerAlaProAspThr-----ArgProAlaProGlySerThr 511
QY 315 ACATTTCCGGCA---GGCTGCTGTTCAACCTCATCTACTCTGCGC-----CAGAAC 362
Db 512 AlaProProAlaHisGlyValThrSerAlaProAspThrArgProAlaProGlySerThr 531
QY 363 GCGGCGTGGCGCTAGCCCTGCTGTTGGACGACAAACACACGCGGGTTGGACGATCTCC 422
Db 532 AlaProProAlaHisGlyValThrSerAlaProAspThrArgProAlaProGlySerThr 551
QY 423 TCCTCGCCCTCG-----ACGCCATCCCAATATCGAAGTGCCTGTTCAACCCCT 473
Db 552 AlaProProAlaHisGlyValThrSerAlaProAspThrArgProAlaProGlySerThr 571
QY 474 TCGTCTACGCAATGGCGCGCACTCGCT-----ACC 506
Db 572 AlaProProAlaHisGlyValThrSerAlaProAspThrArgProAlaProGlySerThr 591
QY 507 TGACCGACTTCCCGCCTCAACCCCGCATGCAACAAATTCCTTTACCGCGCAACCC 566
Db 592 AlaProProAlaHisGlyValThrSerAlaProAspThrArgProAlaProGlySerThr 611
QY 567 GCGCACCATCTCGCGGACCATATCGGCGCAATCTTCAAGTCGCTGAGGACA 626
Db 612 AlaPro-----AlaHisGlyValThrSerAlaProAspThrArgProAla 627
QY 627 CGGTTTTGCGCGACCTGGACATCTCGCCACCGCGCGCTCGTCGCGCAAGTATCGCAG 686
Db 628 ProGlySer-----ThrAlaProProAlaHisGlyValThrSerAlaProAsp 643
QY 687 ACTTCGACCGCTACTGGGCAAGCCATTCGCGC----- 719
Db 644 ThrArgProAlaProGlySerThrAlaProProAlaHisGlyValThrSerAlaProAsp 663
QY 720 ACAACGCGCACGCGCATCTCCGACGGCGCAACATCGGCAAGGGTCTTCAAGCACTCGGAT 779
Db 664 ThrArgProAlaProGlySerThrAlaProProAlaHisGlyValThrSerAlaProAsp 683
QY 780 ACA-----ACGACGAAACATCCAGAC 800
```

Db	684	ThrArgProAlaProGlySerThrAlaProProAlaHisGlyValThrSerAlaProAsp	703
Qy	801	ACGGCGTCTCTGC-----GCTACGCGCAAAACCGTCGAAC-----AGTCGCGCCCTCT	845
Db	704	ThrArgProAlaProGlySerThrAlaProProAlaHisGlyValThrSerAlaProAsp	723
Qy	846	ACCAAAAAATACACACGGGACGATGACTGGCAGAGCGTCCAAACCGCGCTGATCAGCG	905
Db	724	ThrArg-----ProAlaProGlySerThrAlaProProAlaHisGlyVal	738
Qy	906	ACACCCCTGCAAAAGGACTGCACCGCAGCCGCCCAACACCGCGATTGCCGGGAGGTGC	965
Db	739	Thr-SerAlaProAspThrArgProAlaProGlySerThrAlaProProAlaHisGlyVal	758
Qy	966	AAGACGGCGCTCAAAACAGCCCGCAAAAAAGCGCTATCTGTTTCACCCCTATTTCGTC	1025
Db	758	lThrSerAlaProAspThrArgPro-----	766
Qy	1026	CAAAATCGGGCAGACGACTGCAGAACTGGTGCAGAGCGGATAGAGCTTACCGTCC	1085
Db	767	-----AlaProGlySerThrAlaProProAlaHisGlyValThr-Ser	780
Qy	1086	TGACCAACTCGCTACAGCGACCG-----ACGTTGCGCGCGTCCATTCGCGTACGTCA	1139
Db	781	AlaProAspThrArgProAlaProGlySerThrAlaProProAlaHisGlyValThrSer	800
Qy	1140	AATACCCGAAACCCGC-----TGCTCAAAAGCGCGCATCAAACTCTACGAGC	1184
Db	801	AlaProAspThrArgProAlaProGlySerThrAlaProProAlaHisGlyValThrSer	820
Qy	1185	TGCNAACCAACCATGCGTCCCGCCACAAAGACAAGGCGTACCGCA-----	1235
Db	821	AlaProAspThrArgProAlaPro-----GlySerThrAlaProProAlaHisGlyVal	838
Qy	1236	-----GCTCCGTAACACCGCTGCATCGCCAAACCTTCATTG	1271
Db	839	ThrSerAlaProAspThrArgProAlaProGlySerThrAlaProProAlaHisGlyVal	858
Qy	1272	TGACGCGCAACGCATCTTCATCGGCTCATTCACACTCGACCCCGCTTCGCGACGGC--	1328
Db	859	ThrSerAlaProAspThrArgProAlaProGlySerThrAlaProProAlaHisGlyVal	878
Qy	1329	---TCATACCG-----AAATGGCGCTGCTCATCGAAAGCCCAAAATCGCAGACAGA	1379
Db	879	ThrSerAlaProAspThrArgProAlaProGlySerThrAlaProProAlaHisGlyVal	898
Qy	1380	TGGAGCGCACCTCGCGGATACACACCCGAAATACGCTACCGCGTTTACCCCTCGACAAAC	1439
Db	899	ThrSerAlaProAspThrArgProAlaProGlySer---ThrAlaProProAlaHisGly	917
Qy	1440	ACAACCGCCTGCAATGGCAGCATCCCG-----CCACCCGAAAAACCT	1481
Db	918	ValThrSerAlaProAspThrArgProAlaProGlySerThrAlaProProAlaHisGly	937
Qy	1482	ACCGGAACGAAACCG-----AGCCAAACTTTGGAAACCGCATCGCG	1523
Db	938	ValThrSerAlaProAspThrArgProAlaProGlySerThrAlaProProAlaHisGly	953

RESULT 36

```

RES001 36
US-11-079-463-7814
; Sequence 7814, Application US/11079463
; Publication No. US20060073161A1
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO BACTEROIDES FR
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: PATH00-03DIV2
; CURRENT APPLICATION NUMBER: US/11/079,463
; CURRENT FILING DATE: 2005-03-14
; PRIOR APPLICATION NUMBER: US 60/128,705
; PRIOR FILING DATE: 1999-04-09
; PRIOR APPLICATION NUMBER: US 09/540,209

```

```
QY 826 CGTCGACAGTCGCCCTTACCAAAAAATACAGACGGGACGCATCGACTGGCAGAGCGT 885
Db 259 -----LysArg 260
QY 886 CCAAAACCGCGTATCAGCAGCACCCCTGCRAAAGAGTCTGACCGCGACCGCAACCC 945
Db 261 -AsnSerArgMetLeuSerHisThrTyAlaMet----- 271
QY 946 GCCGATTGCGGGAGGCTGCAAGACGCGCTCAACAGCCCGCAAAAGCGTCTATCTGGT 1005
Db 272 -----SerIleTyAlaAlaGlnHisAspValArgIleVa 283
QY 1006 TTCACCTATTTCGTCCTACAAAATCCGGGCACAGACGACTGGCAAAACCTGTCGACGA 1065
Db 283 -AsnProTyPheValProThrSerSerIleArgLysAlaLeuLysArgAlaLeuAsnAr 303
QY 1066 CGGCATAGACGTTACCGTCTGACCAACTCGCTACAGGCGCGGACGCTTGCCGCGTCCA 1125
Db 303 gGlyThrLysValGluIleMetIleSerSer-----LysSerAspIleProPheThrPr 321
QY 1126 TTCGGGTACGTCAAATACGAAACCGCTGCTCAAAGCCGGCCTCAAACTCTACGAGCT 1185
Db 321 oAspAlaSerLeuTyAlaValGlnLysLeuMetLysLysGlyAlaMetIleTyLeuTy 341
QY 1186 GCAACCCACCATGGCTGCCCGCCACAAAAGACAAAGCCCTGACCGGAGCTCCGTAAAC 1245
Db 341 rAsnGlyGlyPheHis----- 346
QY 1246 CAGCTGCATGCAAAACCTTCATTGTGGACGGCAACGCATCTTCATCGGCTCATTTCAA 1305
Db 347 -----HisSerLysIleMetMetValAspAspLeuPheCysThrValGlyThrAlaAs 364
QY 1306 CCTGACCCCGCTTCGCGACGCTCAATACCGAAATGGCGCTGCTCATCGAAAGCCCAA 1365
Db 364 nLeuAsnSerArgSerLeuArgTyAspTyArgLysThrAsnAlaPheIlePheAspLysAs 384
QY 1366 ATCCGACAGACGATGGAGCGCACCCCTC---GCCGAT-----ACCAC 1404
Db 384 pileThrGlnGlnLeuAsnAspValPheGluAlaAspMetLeuHisCysThrArgLeuTh 404
QY 1405 ACCCGAA 1411
Db 404 rProGlu 406

RESULT 37
US-11-096-568A-20771
; Sequence 20771, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nikolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; TITLE OF INVENTION: Theryb
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 20771
; LENGTH: 413
; TYPE: PRT
; ORGANISM: Zea mays subsp. mays
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)..(413)
; OTHER INFORMATION: Ceres Seq. ID no. 12387083
US-11-096-568A-20771

Alignment Scores:
Pred. No.: 7.6e-08 Length: 413
Score: 229.50 Matches: 117
Percent Similarity: 40.2% Conservative: 49
Best Local Similarity: 28.3% Mismatches: 153
Query Match: 8.0% Indels: 94
```

```
DB: 11 Gaps: 20
US-10-665-990A-13 (1-1561) x US-11-096-568A-20771 (1-413)
QY 154 CCTCTGGACAAATCTCTGCAAAATCCGGCACACCCCTCATAAACACGGGCTATCGGACAT 213
Db 18 ProProArgSerAsnGlnSerGlnSerArgProSerGlnArgArg----- 33
QY 214 CTACTGCTCGACGACCCCGACGAAAGCCCTTGC-----CGCCCGCGCGCCCTTAT 264
Db 34 -----AlaArgGluAsnProHisSerProIleProThrProGlnProPheProPro 51
QY 265 CGAATCTGCCAAACACAG-----CCTCGATTGCAATACTACAT 303
Db 52 ArgLeuLeuGlnIleGlnSerSerProSerAlaAsnProProArgLeuGlnIle----- 69
QY 304 TTGGCGCAACGACATTTCCGGCAGCTGCTGTCAACCTCATGTACCTTGGCGGAGAACG 363
Db 70 -----ProGlyArgGlnAlaAlaAsnGlnProTrpSerAlaSerArgArgSer 85
QY 364 -----CGCGGTGCGGTACGCTGCTGTGGACGACGACAAACACCGCGG 408
Db 86 CysSerSerTrpAlaSerSerAlaSerThrAlaSerProArgArgSerSerGluAlaArg 105
QY 409 GTTGCAGCATCT-----CCTGCTCG-----CCTCGA----- 435
Db 106 ThrAlaArgSerSerArgAlaSerArgProProSerArgCysProSerProArgPheSer 125
QY 436 -----CAGCCATCCCAATATCGAAGTCGCTGTTCACCCCTTCCTCTACGCAA 486
Db 126 SerCysThrSerProSerArgThrArgSerSerProThr-SerThrProThrArgPr 145
QY 487 ATGGCGCGCATCTCGGTACCTGACGACTCCCGGCTCAACCGCGCATGCGACAA 546
Db 145 oTrp-ThrHisPro-----ProArgAlaProThrArgSerAlaAlaCysCys----- 160
QY 547 ATCTTTTACCGCCGACACCGCGCACCATACTCGGCGGACGCAATATCGGCGACGAATA 606
Db 161 --ProAlaSerProProThrGlySerProCysSerSerSerSerSerProThrSerSerA 180
QY 607 CTTCAAGTCGTGAGGACA-----CCGTTTTGCGCGACCTGGACATCCT 651
Db 180 rGSerSerCysSerArgSerSerProSerProProProCysSerProSerProSerT 200
QY 652 CGCCACCGCGAGCTGCTGGGGAAGTATCGACGACTCG-----ACCGCTACTG 702
Db 200 hrProProSerAlaThrArgSerProSerArgAlaCysSerProSerCysProAlaSerG 220
QY 703 GCGAAGCCATTCCG-----CCCAACGCGCAC 729
Db 220 lyGlyGlySerProProArgSerSerProProSerArgSerSerSerProThrThrProS 240
QY 730 GC-----GCATCATCCGACGCGCAACATCGGCAAGGCTCTTCAAGCACTCGGATACAA 783
Db 240 erProCysSerSerSerSerAlaSerSerProArg----- 252
QY 784 CGACGAAATCATCAGACACGCGCTCTCGCTACCGCGAAACCGCTGCAACAGT-----CGCC 840
Db 253 --ThrAlaProAlaSerArgAlaCysSer-----ProSerSerProSerP 269
QY 841 CCTCTACCAAAAAATACAGCGGACGATCGACTGGCAGAGCGTCCAAACCCGCTGAT 900
Db 269 roThrCysSerAlaSerSerThrSerAlaSerSerGlyThrSer-Pro----- 284
QY 901 CAGCGACACCCCTGCAAAAGGACTCGACCGCGACCGCGCAACCCCGGATTCGCGGAG 960
Db 285 AlaSerSerProCysSerArgThrThrArgAlaSerProProCysAlaArgAlaArgThr 304
QY 961 GCTGCAAGACGCGCTCAACACCGCGAAAAAGCGCTCTATCTGGTTTACCCCTATTCTGT 1020
Db 305 SerSerGlyGlySerSerProArgValLeuSerSerArgSerSerArgSerSerSer 324
QY 1021 CCTTACAAAAATCCGGCACAGCGCACTGGCAAAACTGTGTGAGGACGCGCATAGAGTTAC 1080
```

```
Db 325 ProProSerSerSerAlaSerAlaProGlySerSerGlyAspProAlaProArgGln 344
QY 1081 CGTCTGACCACTCGCTACAGCGCACCAGCGT-----TGGCGCGGTCCA 1125
Db 345 GlySerCysSerAlaSerCysSerProArgSerAlaSerSerCysSerArgSer 364
QY 1126 TTCCGGCTACGTCAAATACCGAAAAACCGCTGCTCAAAGCCGCGCATCAAACTCTACGAGCT 1185
Db 365 SerArgArg-----ProSerSerThrSerCysAlaArgAlaThrThrArgAla 381
QY 1186 GCAACCAACACCGTCCCGCCGCCACAAAGA 1218
Db 382 SerThrArgProThrSerProThrThrSerArg 392

RESULT 39
US-11-074-176-10
; Sequence 10, Application US/11074176
; Publication No. US20050250135A1
; GENERAL INFORMATION:
; APPLICANT: Klaenhammer, Todd R.
; APPLICANT: Russell, William M.
; APPLICANT: Altermann, Eric
; APPLICANT: McAuliffe, Olivia
; APPLICANT: Peril, Andrea Azcarate
; TITLE OF INVENTION: Nucleic Acid Sequences Encoding
; TITLE OF INVENTION: Stress-Related Proteins and Uses Therefore
; FILE REFERENCE: 5051-694
; CURRENT APPLICATION NUMBER: US/11/074,176
; CURRENT FILING DATE: 2005-03-07
; PRIOR APPLICATION NUMBER: 60/551,161
; PRIOR FILING DATE: 2004-03-08
; NUMBER OF SEQ ID NOS: 381
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 485
; TYPE: PRT
; ORGANISM: Lactobacillus acidophilus
US-11-074-176-10

Alignment Scores:
Pred. No.: 1.05e-07 Length: 485
Score: 227.50 Matches: 99
Percent Similarity: 37.5% Conservative: 65
Best Local Similarity: 22.7% Mismatches: 172
Query Match: 8.0% Indels: 101
DB: 11 Gaps: 16

US-10-665-990A-13 (1-1561) x US-11-074-176-10 (1-485)
QY 149 CCTGTCTCTGGCAACATCTGCAATCGGCGACACCCCTCATACAAAGCGGCTATCC 208
Db 111 ProValThrLysAsnAsnValLysLeu---TyrThrGluGlyAsnThrPhePheHis 129
QY 209 GACATCTACTGCTCGACGACCCCGACGAGCCCTTGGCGCCGCGCGCTTATCGAA 268
Db 130 Asp-----MetLeuLysAsp-----IleVal 136
QY 269 TCTGCGCAACACAGCTCGATTGCAATCTACATATTGGCGCAACGACATTTCCGGCAGG 328
Db 137 ArgAlaLysGluThrIleAsnIleGluPheThrPheThrPheThrPheThrPheThr 156
QY 329 CTGCTGTTCAACCTCATGCTACCTTTGGCGAGAACCGCGCGGCGTACGCTGCTGTG 388
Db 157 ArgValLeuGlnLeuLeuIleLysLysAlaGlnGlnGlyValLysValArgValVal 176
QY 389 GACGACAAACACGCGCGG-----TTGGACGATCTCTGCTCGCCCTC 433
Db 177 AspAlaTrpGlySerMetGlyAlaThrLysAlaThrPheAspGlnLeuArgLysAlaGly 196
QY 434 GACAGCCATCCCAATATGAAGTCGGCTGTTCACCCCTTGTTCACGCAATATGCGGC 493
Db 197 GlyLys-----ValLeuProPheIleThrSerArgAsnMet 208
```

```
QY 494 GCACTCGGCTACCTGACCGACTTCCCGCGCTCAACCCCGCATGCGACAAATCCTTT 553
Db 209 IleThrArgTyr-----ArgIleAsnTyrHisLeuHisArgLysIleVal 223
QY 554 ACCGCCGACAAACCGCGCACCATACTCGCGCGAGCGCAATATCGGCGAGCAATATCTC 610
Db 224 ValIleAspGlyLysIleSerTrpThrGlyGlyPheAsnIleGlyAspGlnTyrLeuGly 243
QY 611 -----AAAGTCGGTGAGGACACCGTTTTTCCCGCACCTGGGACATCTCCGCCACCGC 661
Db 244 ArgLysLysLysPheGlyHis-----TrpArgAspSerGlnValArgIleValGly 260
QY 662 AGCGTCTGCGCGAAGTATCGACGACTTCGACCGCTACTGGGCAAGCCATCTCGCCAC 721
Db 261 SerAlaSerLeuLeuLeuGlnLysPheValMetAspTrp---AsnAlaSerIleAsn 279
QY 722 AACGCCACGCGCATCATCCGCGCGGCAACATCGCGCAAGGCTCTTCAAGCATCGGATAC 781
Db 280 AsnAspAspGluIleIleArg-----Phe 287
QY 782 AACGACGAAACATCCAGACAGCGCTCTCTGGCTACCGGAAACCGTCGAAACAGTCGCC 841
Db 288 Asn-----SerThr 290
QY 842 CTCTACCAAAAAATACAGACGGGAGCGATCGACTGGCGAGCGTCCAAACCGCGCTGATC 901
Db 291 LeuPheProAspLeuAspGluLysAsnIleHisGlnAspAspValAlaThrGlnIle 310
QY 902 AGCGACACCCCTGCAAAAGGACTCGACCGCGACCGCGCAACCGCGATTCGCGGGAGG 961
Db 311 SerAspGlyPro-----AspArgTyrThrSerTyrMetArgAsnGly 324
QY 962 CTGCAAGACGGCTCAAACAGCCCGAAAAAGCGTCTATCTGTTTCCACCTATTTCGTC 1021
Db 325 MetMetArgLeuMetLeuLeuAlaArgAsnArgLeuTrpValGlnThrProTyrLeu 344
QY 1022 CCTACAAAATCCGCGCACAGCGACTGGCAAAACTGGTGCGAGGACGATAGAGCTTACC 1081
Db 345 ProAspAlaValPheAlaThrTrpGlnThrIleAlaMetSerGlyValAspValArg 364
QY 1082 GTCTGTACCAACTCGCTACAGCGACCGAGTTTGGCCCGCTCCATTCCCGCTACGTCAA 1141
Db 365 IleMetIleProCys-----LysProAspHisProPheIleTyrArgAlaThrGlnTrp 382
QY 1142 TACGAAAACCGCTGCTCAAAGCGGATCAAACTCTACGAGTGCACCCCAACCATGCC 1201
Db 383 TyrAlaAsnGluLeuThrArgPheGlyValLysIleTyrIleTyrGluAspGlyPhe 401
QY 1202 GTCCCGCCACAAAAGACAAAGGCTGACCGCGAGCTCCGTAACCGCTCGCATGCCAAA 1261
Db 402 -----LeuHisAlaLys 405
QY 1262 ACCTTCATTGTGACGCGCAAAACGATCTTCATCGGCTCATTCACCTCGACCCCGTTC 1321
Db 406 ThrThrIleIleAspAsnPheSerSerValGlySerMetAsnGlnAspTyrArgSer 425
QY 1322 GCAGGGCTCAATACCGAAATGGGCTGCTCATCGAA-----ACGCCCAAAATC 1369
Db 426 TyrSerLeuAsnPheGluAspAsnAlaIlePheTyrAspLysAsnPheAsnLysMet 445
QY 1370 GCAGACAGATGAGCGCACCCCTCGCGGATACC-----ACACCCGAA 1411
Db 446 AlaGluAlaPheGluAspMetLysLysSerHisLeuLeuThrProGlu 462
```

RESULT 39

```
US-11-072-512-3385
; Sequence 3385, Application US/11072512
; Publication No. US20060029945A1
; GENERAL INFORMATION:
; APPLICANT: ISOGAI, TAKAO
; APPLICANT: SUGIYAMA, TOMOYASU
; APPLICANT: OTSUKI, TETSUJI
```

```
; APPLICANT: WAKAMATSU, AI
; APPLICANT: SATO, HIROYUKI
; APPLICANT: ISHII, SHIZUKO
; APPLICANT: YAMAMOTO, JUN-ICHI
; APPLICANT: ISONO, YUUKO
; APPLICANT: HIO, YURI
; APPLICANT: OTSUKA, KAORU
; APPLICANT: NAGAI, KEIICHI
; APPLICANT: IRIE, RYOTARO
; APPLICANT: TAMECHIKA, ICHIRO
; APPLICANT: SEKI, NAOHIKO
; APPLICANT: YOSHIKAWA, TSUTOMU
; APPLICANT: OTSUKA, MOTOTYUKI
; APPLICANT: NAGAHARI, KENJI
; APPLICANT: MASUHO, YASUHIKO
; TITLE OF INVENTION: Novel full length cDNA
; FILE REFERENCE: 084335-0191
; CURRENT APPLICATION NUMBER: US/11/072,512
; CURRENT FILING DATE: 2005-03-07
; PRIOR APPLICATION NUMBER: US 60/350,978
; PRIOR FILING DATE: 2002-01-25
; PRIOR APPLICATION NUMBER: JP 2001-379298
; PRIOR FILING DATE: 2001-11-05
; NUMBER OF SEQ ID NOS: 4096
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3385
; LENGTH: 715
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-072-512-3385

Alignment Scores:
Pred. No.: 1,11e-07 Length: 715
Score: 227.50 Matches: 156
Percent Similarity: 33.9% Conservative: 75
Best Local Similarity: 22.9% Mismatches: 229
Query Match: 8.0% Indels: 222
DB: 11 Gaps: 31

US-10-665-990A-13 (1-1561) x US-11-072-512-3385 (1-715)

QY 34 CCCGATGAACACGACGCTCATTTCCCTTTTATGCTCCTCTCTGTTGTCATGTTCTTC 93
Db ||| : : : : : ||| ||| |||
QY 4 ProLysArgSerSerLysProSerMetSerLeuAlaProSerGlySerSerMet----- 21
Db ||| : : : : : ||| ||| |||
QY 94 ATGGTTGCCGCCACTGGAAGAACGACGGAAGAGCCG-----TCATTTCAATACTCCAA 147
Db ||| : : : : : ||| ||| |||
QY 22 -----ProThr-----AlaAspProLysProProAlaSerLeuLysSerThrLys 36
Db ||| : : : : : ||| ||| |||
QY 148 ACCTGTCCTCTGGACAACAT-----CCTGCAAAATCCGGCA----- 183
Db ||| : : : : : ||| ||| |||
QY 37 SerAlaThrProAsnArgSerLeuValProThrLysProAlaThrSerArgAsnSerVal 56
Db ||| : : : : : ||| ||| |||
QY 184 ---CACCCCTCATAAACACGGGCTATCCGACATCTACCTGCTCGACGCCGCCACGAAAGC 240
Db ||| : : : : : ||| ||| |||
QY 57 MetSerProSerSerLysSerThrLysSerThrSerThrLysArgAlaProSerAsn 76
Db ||| : : : : : ||| ||| |||
QY 241 CTTGCGCGCCGCGCCCTTTATCGAATC-----TGCCGAACACAGCCTCGATTT 291
Db ||| : : : : : ||| ||| |||
QY 77 -----ArgProSerSerArgSerArgValArgSerLysAlaArgThrProSerArgVal 94
Db ||| : : : : : ||| ||| |||
QY 292 GCAATACATATTCGGCAACGACATTTCCGGCAGGCTGCTGTTCAACCTCATGCTACT 351
Db ||| : : : : : ||| ||| |||
QY 95 Ser-----ThrAspThrArgThrSerLysAlaSerLysAlaSerAspValArg 110
Db ||| : : : : : ||| ||| |||
QY 352 TCCCGAGAACCGCG-----CGTGGCGGTACGCTGCTGTTGGACGACAAACAA 399
Db ||| : : : : : ||| ||| |||
QY 111 CysHisGlnArgArgGlyThrHisSerArgGlyArgThrPro-----GlyArgGly 128
Db ||| : : : : : ||| ||| |||
QY 400 CACGCGCGGTTGGACGATCTCCTGCTCGCCTCGACAGCCATCCCAATATCGAAGTCG 459
Db ||| : : : : : ||| ||| |||
QY 129 SerArgSerLysArgSerProSerArgAlaSerThrProGlyArgIleArgThr--- 147
Db ||| : : : : : ||| ||| |||

460 CTGTTCAACCCCTTCGTCCTACGCAAAATGCGCGGCACTCGGCTACCT----- 507
148 -----HisGlyAlaArgProGlyMetAlaSerValArgThrProThrSerSerGlnGln 165
507 ----- 507
166 LysGlySerArgGlyLysSerTyrGlyArgProArgThrSerAsnArgGluArgSerAsp 185
508 -----GACCGACTTCCCGGCC----- 524
186 SerGlnProArgAsnLeuSerLysLysSerTyrArgProGlyGlySerGlyIleGly 205
525 -----TCAACCCGCGCATGCACACAAATCCTTTACCG 557
206 ArgSerSerGluLeuAlaValThrProSerThrAlaLysCysGlnThr----- 221
558 CCGACAAACCGCGGCCACCATCTCGCGGACCAATATCGCGGACGAATACTTCAAAGTCG 617
222 ProThrGlyIleProSerLysGluLysSerAspAsnProSerProSerSerSerArgLys 241
618 GTGAGGACA-----CCGTTTTCGCCGACCTGG 644
242 VallysSerTyrGlyGlnMetIleProSerArgGluLysSerTyrSerProThrGlu 261
645 ACATCCTCG---CCACCGGACGCTCGTCGGGAAGTATCGCAGCACTTCCACCGCTACT 701
262 MetSerSerArgValLysSerTyrAsnGlnAlaSerThrArgSerArgProGlnSerHis 281
702 GGGCAAGCCATTCCGCCACACCAACCGCATCTCCGCGGCGGCAACATCGGCAAGG 761
282 SerGlnSerArgSerProArgSerArgSerGlySerGlnLysArgThrHisSerArg 301
762 GTCTTCAAGCACTCG----- 776
302 ValArgSerHisSerTrpLysArgAsnHisSerArgAlaArgSerArgThrArgLysGly 321
777 -----GATACAACGACGAAACATCCACACGCGCTCTCGCTACCGCAACCGCTCG 830
322 IleLeuSerGlnMetGlyArgHisSerGlnSerArgSerHisSerLysGlyLysSerGln 341
831 AAC----- 833
342 AsnGlnSerArgThrProArgArgGlyArgSerHisAsnTrpSerArgAsnProSerLys 361
834 -----AGTCGCCCTCTACCAAAAAAATACACGCGGACGCTCGACT 875
362 GluArgSerHisSerHisSerArgSerSerLysGlu-----ArgAsp-HisArgGly 379
876 GGCAGAGCGTCCA-----AACCGCGCTGATCAG-- 903
379 ySerSerSerProArgLysGluSerGlyArgSerGlnSerGlySerProAsnLysGlnAr 399
904 -----CGACACCCCTGCAAAAGGACTCGACCGCGACCCGCGCAACCGC 947
399 gAspHisSerArgSerArg-SerProAsnLysAlaArgAspArgSerArgSerArgSerP 419
948 CGATTGCGGAGGCTGCAAGACCGCTCAACAGCCGCAAAAGCGTCTACTCTGGTTT 1007
419 roTyrLysAlaArg-AspArgSerArgSerArgSerProAsnLysAlaArgAspCysSer 438
1008 CACCTATTTCGTCCTACAAAATCCGGCAGACGCACTGGCAAACT-----GGTGC 1061
439 ArgSerArgSer-ProTyrLysAlaArgAspArgSerArgSerArgSerArgSerArgSe 478
1062 AGGACGGCATAGAGCTTACCGTCTCTGACCAACTCGCT----- 1098
458 aArgAspHisSerArgSerArgSerProAsnLysAlaArgAspArgSerArgSerArgSe 478
1099 -----ACAGCGACCG 1109
478 rProSerLysGluArgAspHisSerGlnLeuGlySerProSerLysGluArgAspHisAr 498
1110 ACGTTGCGCGCTCCATTCCGCTACGTCAAAATACCGAAAAACCGCTGCTCAAGCGCGCA 1169
```


Qy 1418 TACCGGTTACCGCTCGACAAACACACACCGCGCTGCAATGG 1456
Db 377 LysArgProValMetGlnArg-----ProValGluTrp 387

Search completed: May 2, 2006, 06:05:44
Job time : 124.5 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: May 2, 2006, 05:02:04 ; Search time 76 Seconds
(without alignments)
2898.236 Million cell updates/sec

Title: US-10-665-990A-13
Perfect score: 2852
Sequence: 1 caaatacaggaatgcgt.....tgccatcgaaggtttatta 1561

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 4332886

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:
-MODEL=frame+ n2p.model -DEV=xlp
-Q=/abss/ABSSWEB.spool/US10665990/runat_01052006_111941_9494/app.query.fasta_1
-DB=UniProt -QWTF=fastan -SUFFIX=rup -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=BLOSUM62 -TRANS=human40.cdi -LIST=45
-DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=40 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZES=500 -MINLEN=0 -MAXLEN=200000000 -HOST=abs02p
-USER=US10665990 @CGN 1 1 466 @runat_01052006_111941_9494 -NCPU=6 -ICPU=3
-NO MMAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DRV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : UniProt 05.80.*
1: uniprot.sprot.*
2: uniprot.trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2689	94.3	525	Q6W764_NEIGO	Q6W764 neisseria 9
2	2681	94.0	525	Q5F887_NEIG1	Q5F887 neisseria 9
3	2630	92.2	525	Q9JTT1_NEIMA	Q9JTT1 neisseria m
4	2573	90.2	508	Q9JYU0_NEIMB	Q9JYU0 neisseria m
5	1182	41.4	524	Q7VUJ6_BORPE	Q7VUJ6 bordetella
6	1181	41.4	492	Q7WP63_BORBR	Q7WP63 bordetella
7	1104.5	38.7	493	Q8FIS6_ECOL6	Q8FIS6 escherichia
8	1102.5	38.7	493	1_YNDC_ECOLI	P75919 escherichia
9	1102.5	38.7	493	Q3RUS5_SHIFL	Q3RUS5 shigella fl
10	1100.5	38.6	493	Q7AFAB_ECO57	Q7AFAB escherichia
11	1098	38.5	528	Q57QL0_SALCH	Q57QL0 salmonella
12	1093.5	38.3	493	Q8X917_ECO57	Q8X917 escherichia
13	1092.5	38.3	514	Q9CKM5_PASMU	Q9CKM5 pasteurella
14	1089.5	38.2	494	Q8ZQ28_SALTY	Q8ZQ28 salmonella
15	1089.5	38.2	495	Q8Z7M0_SALTI	Q8Z7M0 salmonella
16	1089.5	38.2	495	Q5PGY9_SALPA	Q5PGY9 salmonella

17	940.5	33.0	519	2	Q6F8A6_AC1AD	Q6F8A6 acinetobact
18	881.5	30.9	516	2	Q4LR64_9BURK	Q4LR64 burkholderi
19	881	30.9	505	2	Q87J92_VIBPA	Q87J92 vibrio para
20	872.5	30.6	501	2	Q8D641_VIBVU	Q8D641 vibrio vuln
21	869.5	30.5	524	2	Q8D861_VIBVU	Q8D861 vibrio vuln
22	863.5	30.3	554	2	Q7MMB3_VIBVY	Q7MMB3 vibrio vuln
23	862.5	30.2	501	2	Q8D800_VIBVU	Q8D800 vibrio vuln
24	852	29.9	570	2	Q4FTL4_9GAMM	Q4FTL4 psychrobact
25	850.5	29.8	520	2	Q8PIM4_XANAC	Q8PIM4 xanthomonas
26	849.5	29.8	530	2	Q6MQL7_BDEBA	Q6MQL7 bdellovibri
27	846	29.7	464	2	Q5QUY8_IDILO	Q5QUY8 idiomarina
28	843	29.6	520	2	Q4UWU7_XANCP	Q4UWU7 xanthomonas
29	843	29.6	520	2	Q8P7A5_XANCP	Q8P7A5 xanthomonas
30	828	29.0	523	2	Q6G185_BAROU	Q6G185 bartonella
31	825	28.9	523	2	Q6G4X0_BARHE	Q6G4X0 bartonella
32	812.5	28.5	466	2	Q98N23_RHILO	Q98N23 rhizobium l
33	807	28.3	518	2	Q8UEX3_AGRTS	Q8UEX3 agrobacteri
34	804.5	28.2	466	2	Q984B3_RHILO	Q984B3 rhizobium l
35	794	27.8	520	2	Q6F853_AC1AD	Q6F853 acinetobact
36	792	27.8	529	2	Q9HTP4_PSEAE	Q9HTP4 pseudomonas
37	774	27.1	550	2	Q62E43_BURVA	Q62E43 burkholderi
38	773	27.1	517	2	Q88CA5_PSEPK	Q88CA5 pseudomonas
39	772.5	27.1	540	2	Q4LJD9_9BURK	Q4LJD9 burkholderi
40	769	27.0	542	2	Q88BC2_PSESM	Q88BC2 pseudomonas
41	766	26.9	550	2	Q63J33_BURPS	Q63J33 burkholderi
42	762.5	26.7	510	2	Q6LII3_PHOPR	Q6LII3 photobacter
43	755.5	26.5	526	2	Q4ZZX0_PSESY	Q4ZZX0 pseudomonas
44	751	26.3	515	2	Q5E616_VIBFI	Q5E616 vibrio fisc
45	750	26.3	521	2	Q92QQ3_RHIME	Q92QQ3 rhizobium m

ALIGNMENTS

RESULT 1
Q6W764_NEIGO
ID Q6W764_NEIGO PRELIMINARY; PRT; 525 AA.
AC Q6W764;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Phospholipase D (fragment).
GN Name=pld;
OS Neisseria gonorrhoeae.
OC Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC Neisseriaceae; Neisseria.
OX NCBI_TaxID=485;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=1291;
RX MEDLINE=22935225; PubMed=14573659;
RX DOI=10.1128/IAI.71.11.6381-6391.2003;
RA Edwards J.L., Entz D.D., Apicella M.A.;
RT "Gonococcal phospholipase d modulates the expression and function of
RT complement receptor 3 in primary cervical epithelial cells.";
RL Infect. Immun. 71:6381-6391(2003).
DR EMBL; AY307929; AAQ77232.1; -; Genomic DNA.
DR GO; GO:0003824; F:catalytic activity; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR001736; PLD.
DR Pfam; PF00614; PLDC; 2.
DR SMART; SM00155; PLDC; 2.
DR PROSITE; PS50035; PLD; 2.
FT NON TER 525 525
SQ SEQUENCE 525 AA; 59224 MW; 4B626F5264E076D1 CRC64;

Alignment Scores:
Pred. No.: 6.33e-142 Length: 525
Score: 2689.00 Matches: 520
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 94.3% Indels: 0
DB: 2 Gaps: 0

[illegible]

Db	366	ValLeuThrAsnSerLeuGlnAlaThrAspValAlaAlaValHisSerGlyTyrValLys	385		
QY	1142	TACCGAAAAACCGCTCTCAAGCGCGGCATCAAACTCTACGAGCTGCAACCAACCATGCC	1201		
Db	386	TyrArgLysProLeuLeuLysAlaGlyLeuLysLeuTyrGluLeuGlnProAsnHisAla	405		
QY	1202	GTCCCCCGCCACAAAGACAAAGGCGCTACCGGAGCTCCGTAACGACGCTGCATGCCAAA	1261		
Db	406	ValProAlaThrLysAspLysGlyLeuThrGlySerSerValThrSerLeuHisAlaLys	425		
QY	1262	ACCTTCATCTGGAGCGCAACGACGCTTCATCGGCTCATCAACCTCGAGCCCCCGTTCC	1321		
Db	426	ThrPheIleValAspGlyLysArgIlePheIleGlySerPheAsnLeuAspProArgSer	445		
QY	1322	GCACGGCTCAATACCGAAATGGGGCTCGTCATCGAAAGCCCCAAATCGCAGAACAGATG	1381		
Db	446	AlaArgLeuAsnThrGluMetGlyValIleGluSerProLysIleAlaGluGlnMet	465		
QY	1382	GAGCGCACCTTCGCCGATACACACCCGAATACGCTTACCGCGTTACCTTCGCACAAACAC	1441		
Db	466	GluArgThrLeuAlaAspThrThrProGluTyrAlaTyrArgValThrLeuAspLysHis	485		
QY	1442	AACCGCTGCATTCGACGAGTATCCCGCCACCCGAAAAAACCCTACCGCAACCCGAGCC	1501		
Db	486	AsnArgLeuGlnThrHisAspProAlaThrArgLysThrTyrProAsnGluProGluAla	505		
QY	1502	AAACTTTGGAAACGCATCGCGCGAAAAATCCTATCCCTGCTGCCATCGAAGCTTTATTA	1561		
Db	506	LysLeuThrLysArgIleAlaAlaLysIleLeuSerLeuLeuProIleGluGlyLeuLeu	525		
RESULT 3					
Q9JTT1 NEIMA					
ID	Q9JTT1	NEIMA PRELIMINARY;	PRT; 525 AA.		
AC	Q9JTT1				
DT	01-OCT-2000	(TrEMBLrel. 15, Created)			
DT	01-OCT-2000	(TrEMBLrel. 15, Last sequence update)			
DT	01-MAR-2004	(TrEMBLrel. 26, Last annotation update)			
DE	Phospholipase D-family protein.				
GN	OrderedLocusNames=NMA1646;				
OS	Neisseria meningitidis (serogroup A).				
OC	Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;				
OC	Neisseriaceae; Neisseria.				
OX	NCBI_TaxID=65699;				
RN	[1]				
RP	NUCLEOTIDE SEQUENCE.				
RC	STRAIN=Z491 / Serogroup A / Serotype 4A;				
RX	MEDLINE=20222556; PubMed=10761919; DOI=10.1038/35006655;				
RA	Parkhill J., Achtman M., James K.D., Bentley S.D., Churcher C.M.,				
RA	Klee S.R., Morelli G., Basham D., Brown D., Chillingworth T.,				
RA	Davies R.M., Davis P., Devlin K., Feltwell T., Hamlin N., Holroyd S.,				
RA	Jagels K., Leather S., Moule S., Mungall K.L., Quail M.A.,				
RA	Rajandream M.A., Rutherford K.M., Simmonds M., Skelton J.,				
RA	Whitehead S., Spratt B.G., Barrall B.G.;				
RT	"Complete DNA sequence of a serogroup A strain of Neisseria				
RT	meningitidis Z491.";				
RL	Nature 404:502-506(2000).				
DR	EMBL:	AL162756; CAB84874.1; -; Genomic_DNA.			
DR	PIR:	B81859; B81859.			
DR	GO:	GO:0003824; F:catalytic activity; IEA.			
DR	GO:	GO:0008152; P:metabolism; IEA.			
DR	InterPro:	IPR001736; PLD.			
DR	Pfam:	PF00614; PLDC; 2.			
DR	SMART:	SM00155; PLDC; 2.			
DR	PROSITE:	PS50035; PLD; 2.			
KW	Complete proteome.				
SQ	SEQUENCE 525 AA; 59305 MW; 1CFB7AC5B82F1B02 CRC64;				
Alignment Scores:					
Pred. No.:	1.24e-138	Length:	525		
Score:	2630.00	Matches:	509		
Percent Similarity:	98.7%	Conservative:	7		
Best Local Similarity:	97.9%	Mismatches:	0		
Query Match:	92.2%	Indels:	0		

DB: 2 Gaps: 0

US-10-665-990A-13 (1-1561) x Q9UTTI_NEIMA (1-525)

QY 2 AAAATACAGGCAATCCGCTCTGAAACTATATCCCGATGAAACACGAGCCTCATTTCC 61
|||||
Db 6 LysThrGlnAlaMetProSerGluThrIleSerLeuMetLysThrArgSerLeuIleSer 25
|||||
QY 62 CTTTATGCTCCTCTCTCTGTTCAATGTTCTTCATGTTGCCCGCCACTGGAAGACGGAGC 121
|||||
Db 26 LeuLeuCysLeuLeuLeuCysSerSerTrpLeuProLeuProLeuGluGluArgThr 45
|||||
QY 122 GAAAGCCCTCATTTCAATACTTCCAAACCTGTCCTCTCGGACACATCCTGCAATCCGG 181
|||||
Db 46 GluSerArgHisPheAsnThrSerLysProValArgLeuAspAsnIleLeuGlnIleArg 65
|||||
QY 182 CACACCCCTCATAAACAAAGGGGTATCCGACATCTACCTGCTCGACGACCCCGCAAGGCC 241
|||||
Db 66 HisThrProHisThrAsnGlyLeuSerAspIleTyrLeuLeuAsnAspProHisGluAla 85
|||||
QY 242 CTTGCGCGCGCGCGCTTATCGAATCTGCGGAACACAGCCTCGATTTGCAATPACTAC 301
|||||
Db 86 PheAlaAlaArgAlaAlaLeuIleGluSerAlaGluHisSerLeuAspLeuGlnTyrTyr 105
|||||
QY 302 ATTTGGCGACGACATTTCCGGCAGGCTGCTGTTCAACCTCATGTACTTGGCCGACGAA 361
|||||
Db 106 IleTrpArgAsnAspIleSerGlyArgLeuLeuPheAsnLeuValTyrLeuAlaAlaGlu 125
|||||
QY 362 CCGCGCTGCGGCTACGCTGCTGTGGACGACAAACACGCGCGGCTTGGACGATCTC 421
|||||
Db 126 ArgGlyValArgValArgLeuLeuLeuAspAsnAsnThrArgGlyLeuAspAspLeu 145
|||||
QY 422 CTGCTCGCGCTCGACAGCCATCCCAATATCGAAGTGCCTGTTCAACCCCTCTGCTCTA 481
|||||
Db 146 LeuLeuAlaLeuAspSerHisProAsnIleGluValArgLeuPheAsnProPheValLeu 165
|||||
QY 482 CGCAATGCGCGGCTACGCTACCTGACCGACTTCCCGCGCTCAACGCGCGCATGAC 541
|||||
Db 166 ArgLysTrpArgAlaLeuGlyTyrLeuThrAspPheProArgLeuAsnArgArgMetHis 185
|||||
QY 542 AACAAATCTTTACCGCGCACAAACCGCGCACCATCTCGGGGACGCAATATCGCGCAC 601
|||||
Db 186 AsnLysSerPheThrAlaAspAsnArgAlaThrIleLeuGlyGlyArgAsnIleGlyAsp 205
|||||
QY 602 GAATACTTCAAAGTCGGTGAGACACCGTTCCTGCGCAGCTGGAATCCTCGCCACCGGC 661
|||||
Db 206 GluTyrPheLysValGlyGluAspThrValPheAlaAspLeuAspIleLeuAlaThrGly 225
|||||
QY 662 AGCGTCGTCGGGAAGTATCGCACGCTACGACCGCTACTGGCGAAGCATTCGCGCCAC 721
|||||
Db 226 SerValValGlyGluValSerHisAspPheAspArgTyrTrpAlaSerHisSerAlaHis 245
|||||
QY 722 AACGCGACGCGCATCATCCGACGCGCAACATCGGCAAGGGTCTTCAAGCACTCGGATAC 781
|||||
Db 246 AsnAlaThrArgIleIleArgSerGlyAsnIleGlyLysGlyLeuGlnAlaLeuGlyTyr 265
|||||
QY 782 AACGACGAAACATCCAGACACCGCTCTGCGCTACCGCGAAACCGTCGAAACAGTCGCCC 841
|||||
Db 266 AsnAspGluThrSerArgHisAlaLeuLeuArgTyrArgGluThrValGluGlnSerPro 285
|||||
QY 842 CTTCTACCAAAATACAGCGGACGATCGATCGAGCGGTCCAAACCGCTGATC 901
|||||
Db 286 LeuTyrGlnLysIleGlnThrGlyArgIleAspTrpGlnSerValGlnThrArgLeuIle 305
|||||
QY 902 AGCGACACCCCTGCAAAAGGACTCGACGCGGACCGCGCAACCCCGCATTTGCCGGAGG 961
|||||
Db 306 SerAspAspProAlaLysGlyLeuAspArgAspArgArgLysProIleAlaGlyArg 325
|||||
QY 962 CTGCAAGACGCGCTCAAAACAGCCCGGAAAGCGTCTATCTGGTTTACCCCTATTTCTGTC 1021
|||||
Db 326 LeuGlnAspAlaLeuLysGlnProGluLysSerValTyrLeuValSerProTyrPheVal 345
|||||
QY 1022 CCTCAAAATCCGGCACAGCCACTTGGCAAAACTGGTGCAGGACGGCATAGACGTTACC 1081
|||||

Db 346 ProThrLysSerGlyThrAspAlaLeuAlaLysLeuValGlnAspGlyIleAspValThr 365
|||||
QY 1082 GTCTGTACCAACTCGCTACAGGCGACGAGCTTCCCGCGCTCCATTCGGGCTAGCTCAAA 1141
|||||
Db 366 ValLeuThrAsnSerLeuGlnAlaThrAspValAlaAlaValHisSerGlyTyrValLys 385
|||||
QY 1142 TACCGAAAAACCGCTGCTCAAAAGCGGATCAAACTCTACGAGCTGCAACCAACCATGCC 1201
|||||
Db 386 TyrArgLysProLeuLeuLysAlaGlyIleLysLeuTyrGluLeuGlnProAsnHisAla 405
|||||
QY 1202 GTCCCGCGCCACAAAGAGCGCTGACCGGAGCTCCGTAACACAGCCTGCATGCCAAA 1261
|||||
Db 406 ValProAlaThrLysAspLysGlyLeuThrGlySerSerValThrSerLeuHisAlaLys 425
|||||
QY 1262 ACCTTCATTTGAGCGCAACCGCATCTTCATCGGCTCATTCACCTCGACCCCGTTCC 1321
|||||
Db 426 ThrPheIleValAspGlyLysArgIlePheIleGlySerPheAsnLeuAspProArgSer 445
|||||
QY 1322 GCAGCGCTCAATACCGAAATGGCGCTGCTCATCGAAAGCCCCAAAATCGGAGACAGATG 1381
|||||
Db 446 AlaArgLeuAsnThrGluMetGlyValIleGluSerProLysIleAlaGluGlnMet 465
|||||
QY 1382 GAGCGCACCTCGCGCATACCCACACCGCAATACCGCTACCGCTTACCTCGACAAACAC 1441
|||||
Db 466 GluArgThrLeuAlaAspThrSerProGluTyrAlaTyrArgValThrLeuAspArgHis 485
|||||
QY 1442 AACCGCTGCAATGGCACGATCCCGCCACCCGAAAAACCTACCCGAAACGAAACCCGAGCC 1501
|||||
Db 486 AsnArgLeuGlnTrpHisAspProAlaThrArgLysThrTyrProAsnGluProGluAla 505
|||||
QY 1502 AAACCTTGGAAACGATCGCGCGGAAAAATCTATCTCTGCTGCCATCGAAGGTTTATTA 1561
|||||
Db 506 LysLeuTrpLysArgIleAlaAlaLysIleLeuSerLeuLeuProIleGluSerLeuLeu 525
|||||

RESULT 4

Q9UYU0_NEIMB PRELIMINARY; PRT; 508 AA.

ID Q9UYU0_NEIMB PRELIMINARY; PRT; 508 AA.

AC Q9UYU0;

DT 01-OCT-2000 (TrEMBLrel. 15, Created)

DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)

DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)

DE Cardiolipin synthetase family protein.

GN Ordered locus names=NM1434;

OS Neisseria meningitidis (serogroup B).

OC Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;

OC Neisseriaceae; Neisseria.

OX NCBI_TaxID=491;

OX [1]

RN NUCLEOTIDE SEQUENCE.

RP STRAIN=MC58 / Serogroup B;

RC MEDLINE=20175755; PubMed=10710307; DOI=10.1126/science.287.5459.1809;

RX Tetelin H., Saunders N.J., Heidelberg J.F., Jeffries A.C.,

RA Nelson K.E., Eissen J.A., Ketchum K.A., Hood D.W., Peden J.F.,

RA Dodson R.J., Nelson W.C., Gwinn M.L., DeBoy R.T., Peterson J.D.,

RA Hickey E.K., Haft D.H., Salzberg S.L., White O., Fleischmann R.D.,

RA Dougherty B.A., Mason T.M., Ciecko A., Parksey D.S., Blair E.,

RA Citron H., Clark E.B., Cotton M.D., Utterback T.R., Khouri H.M.,

RA Qin H., Vamathevan J.J., Gill J., Scarlato V., Masignani V., Pizza M.,

RA Grandi G., Sun L., Smith H.O., Fraser C.M., Moxon E.R., Rappuoli R.,

RA Venter J.C.;

RT "Complete genome sequence of Neisseria meningitidis serogroup B strain

RT MC58.";

RL Science 287:1809-1815(2000).

DR EMBL; AE002098; AAF41795.1; -; Genomic_DNA.

DR PIR; B81083; B81083.

DR TIGR; NWB1434; -.

DR GO; GO:0003824; P: catalytic activity; IEA.

DR GO; GO:0008152; P: metabolism; IEA.

DR InterPro; IPR001736; PLD.

DR Pfam; PF00614; PLDC; 2.

DR SMART; SM00155; PLDC; 2.

DR PROSITE; PS50035; PLD; 2.

RA	Parkhill J., Sebahnia M., Preston A., Murphy L.D., Thomson N.R.,	
RA	Harris D.E., Holden M.T.G., Churcher C.M., Bentley S.D., Mungall K.L.,	
RA	Cerdeno-Tarraga A.-M., Temple L., James K.D., Harris B., Quail M.A.,	
RA	Achtman M., Atkin R., Baker S., Basham A., Bason N., Cherevach I.,	
RA	Chillingworth T., Collins M., Cronin A., Davis P., Doggett J.	
RA	Feltwell T., Goble A., Hamlin N., Hauser H., Holroyd S., Jegelis K.,	
RA	Leather S., Moule S., Norberczak H., O'Neil S., Ormond D., Price C.,	
RA	Rabinowitz E., Rutter S., Sanders M., Saunders D., Seeger K.,	
RA	Shaw S., Simmonds M., Skellon J., Squares R., Squares S., Stevens K.,	
RA	Unwin L., Whitehead S., Barrell B.G., Maskell D.J.;	
RT	Comparative analysis of the genome sequences of Bordetella pertussis,	
RT	Bordetella parapertussis and Bordetella bronchiseptica.;	
RL	Nat. Genet. 35:32-40(2003).	
DR	EWEL, BX640439, CAE33321.1; -; Genomic DNA.	
DR	GO: GO:0003824; F: catalytic activity; IEA.	
DR	GO: GO:0008152; P: metabolism; IEA.	
DR	InterPro: IPR001736; PLD.	
DR	Pfam: PF00614; PLDC; 2.	
DR	SMART: SM00155; PLDC; 2.	
DR	PROSITE: PS50035; PLD; 2.	
KW	Complete proteome.	
SQ	SEQUENCE 492 AA; 53309 MW; BC6BFBE8AC34E726 CRC64;	

Alignment Scores:

Pred. No.:	9.63e-58	Length:	492
Score:	1181.00	Matches:	253
Percent Similarity:	63.3%	Conservative:	70
Best Local Similarity:	49.6%	Mismatches:	147
Query Match:	41.4%	Indels:	40
DB:	2	Gaps:	8

US-10-665-990A-13 (1-1561) x Q7WP63 BORBR (1-492)

98	QY	TTGCCCCCACTGAGAACCGGACGAAAGCGTCAATTTCCAATACTTCCAACCTGTCCTC	157
1	MetProProValAspArgThrAlaSerHisAlaLeuAspAlaSer	16	
158	QY	CTGGCAACATCCTCGAAATCGGCACACCCCTCATAAACACGGGCTATCCGACATC	214
17	Db	-----GlnAlaArgAlaThrProLeuGlyGlnGlyIleGlyAspLeuAla	31
215	QY	-----TACCTGTCTCGACGACCCGCCCAAGCGCTTGCC	247
32	Db	AspAlaHisProGlyLeuSerGlyPheHisAlaLeuGlyAspAlaGlnAspAlaPheAla	51
248	QY	GCCCGCGCCCTTATCGAATCTGCCAAACACAGCCTCGATTGTCGAATCTACATTGG	307
52	Db	AlaArgMetLeuLeuAlaArgAlaAlaThrArgThrLeuAspValGlnTyrIleTip	71
308	QY	CGCAACGACATTTCCGGCAGGCTGTGTTCACCTCATGTACCTTGCCGACAGACCGCG	367
72	Db	ArgAsnAspMetThrGlyThrLeuLeuLeuGlnAlaLeuHisAlaAlaAlaGluArgGly	91
368	QY	GTGCGCGTACGCTGCTGTGAGACACAAACACGCGCGGTGGACGATCTCTGCTC	427
92	Db	ValArgValArgLeuLeuLeuAspAspAsnGlyIleSerGlyLeuAspAspAlaLeuAla	111
428	QY	GCCTCTGCAGCGCATCCCAATATCGAAGTGGCGCTGTTCACCCCTTCGCTCTACGCAA	487
112	Db	AlaLeuAspAlaHisProAsnAlaGluValArgLeuPheAsnProPheProThrArgSer	131
488	QY	TGGCGGCACTCGGTACTCTGACCGACTTCCCGCGCTCAACCGCGCATGCAACA	547
132	Db	PhcLysAlaLeuGlyTyrLeuThrAspPheSerArgLeuAsnArgArgMetHisAsnLys	151
548	QY	TCCTTTACGCGCACACCGGCCACCATATCTCGCGGACGCAATATCGCGCAGCAATAC	607
152	Db	SerPheThrValAspAsnGlnAlaThrIleIleGlyGlyArgAsnIleGlyAspGluTyr	171
608	QY	TTCAAGAGCTGGTAGGACACCGTTTTTCGCGACCTGGACATCTCTGCGCACCGGACGCTC	667
172	Db	PheGlyAlaThrAspGlyValLeuPheAlaAspLeuAspValLeuAlaValGlyProVal	191

QY	668	GTCCGCGAAGTATCGCAGCATTCGACCGCTACTGGCGAAGCCATTCCGCCCAACGCC	727		
Db	192	ValGlySerValSerAlaGluPheAspAlaTyrTrpAlaSerGluSerAlaThrProAla	211		
QY	728	ACGCGCATCATCCGACGGCGACATCGGCAAGGTTCTTCAAGCACTCGGATACACGAC	787		
Db	212	GlyProLeuLeuProAla-----ProGlyAlaGlnThrLeuArgAlaLeuAla	227		
QY	788	GAACACATCCAGACAC-----GCGCTCTCGCTACCGCAAAACCGTCCAA	832		
Db	228	GluArgAlaAlaAArgIleGluGlnAspProAlaAlaGlyAspTyrMetSerAlaLeuArg	247		
QY	833	CAGTCGCCCTCTACCAAAAATAACAGCGGACGACATCGACTGGCAGACGCTCCAAACC	892		
Db	248	LeuLeuProPheIleArgGluLeuMetAlaGlyArgLeuProLeuGlnThrPalaProAla	267		
QY	893	CGCCTGATCAGCAGACCCCTGCAAAAGGACTCGACCGCGACCGCGCAACCGCGATT	952		
Db	268	ArgMetValSerAspAspProAlaLysGlyLeuGly-----LysAlaProPro	283		
QY	953	CGCCGG-----AGCTGCAAGACGCGTCAAAACAGCCGCAAAAGCGTCTAT	1000		
Db	284	AlaGlyLeuLeuThrGlnGlnLeuArgAsnIleLeuGlyProArgArgThrLeuAsp	303		
QY	1001	CTGTTTCCACCTATTTTCCTCCCTACAAAAATCCGGCAGCAGACACTGGCAAACTGGT	1060		
Db	304	LeuValSerProTyrPheValProThrGluAlaGlyThrGlnAlaPheAlaLeuAla	323		
QY	1061	CAGACCGCATAGACGTTACGGTCTGTGACCAACTCGTACAGGGGACCGAGTTGGCGCC	1120		
Db	324	ArgGlyGlyAlaGlnValArgValLeuThrAsnAlaLeuGluAlaThrAspValAlaVal	343		
QY	1121	GTCCATTCCCGCTACGTCAAATACCGAAAAACCGCTCTCAAACCGCGCATCAAACTCTAC	1180		
Db	344	ValHisserGlyTyrAlaLysArgGlyAlaLeuLeuGlnAlaGlyValArgLeuTyr	363		
QY	1181	GAGCTGCAACCCCAACCATGCGCTCCCGGCCACAAAAAGACAAAGCGCTGACCGGACGCTC	1240		
Db	364	GluMetArgArgSerTyrAlaGlyProLysProGlnGlyArgGlyArgPheGlySerSer	383		
QY	1241	GTAACGAGCTGCATGCCAAAAACCTTCATTGTGGACGGCAACGCATCTTCATCGGCTCA	1300		
Db	384	GlySerSerLeuHisAlaLysThrPheGlyValAspGlyGluArgValPheIleGlySer	403		
QY	1301	TTCAACCTCCACCCCGTTCGCGACGCTCAATACCGAAATGGCGTGTCTACGAAAGC	1360		
Db	404	PheAsnPheAspProArgSerAlaAsnLeuAsnThrGluLeuGlyPheValIleGluSer	423		
QY	1361	CCCAAAATCGCAACAGATGGAGCGCACCTCCGCGATACCAACCCGAATACGCGCTAC	1420		
Db	424	ProAspMetAlaArgHisIleAlaAlaThrPheAspGlnAspIleProAlaAlaThrTyr	443		
QY	1421	CGGTTTACCTTCGCAAAACACACCCGCTGCAATGG-----CACGATCCCGCCACC	1471		
Db	444	GluValArgLeuAspAspAspGlySerLeuTyrTrpLeuGlnArgAspGlyAlaThr	463		
QY	1472	CGAAAAACCTTACCGCAACGAACCCGAAGCGCAAACTTTGGAAACGCATCGCGCAAAATC	1531		
Db	464	ValArg---HisAspSerGluProGlyValSerLeuThrArgArgPheSerValTrpLeu	482		
QY	1532	CTATCCCTGTCGCCCATCGAAGGTTTATTA	1561		
Db	483	PheSerLeuLeuProLeuGluProLeuLeu	492		
RESULT 7					
Q8FIS6_ECOL6					
ID	Q8FIS6_ECOL6	PRELIMINARY;	PRT; 493 AA.		
AC	Q8FIS6;				
DT	01-MAR-2003	(TrEMBLrel. 23, Created)			
DT	01-MAR-2003	(TrEMBLrel. 23, Last sequence update)			
DT	01-OCT-2003	(TrEMBLrel. 25, Last annotation update)			
DE	Hypothetical protein ymdc.				
GN	Name=ymdc; OrderedLocusNames=c1310;				

RESULT 7

08FIS6 ECOL6

Q08150 ECOL6
ID 08FIS6 ECOL6 PRELIMINARY: PRT: 493 AA.

```
AC Q8FIS6:  
001100 00000 1000000000,  
INT; 193
```

DT 01-MAR-2003 (TrEMBLrel. 23, Created)

DT 01-MAR-2003 (TRENBLrel. 23, Last sequence update)

DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)

DE Hypothetical protein ymdC.

GN Name=yndC; OrderedLocusNames=c1310;


```
QY 1490 GAACCCGAGCCAAACTTTGGAAAGCGCATCGCCGCAAAATACTATCTCTGCTGCGCCATC 1549
D 1490 GUAProAlaThrSerPheTrpIysArgValMetValArgLeuAlaSerIleLeuProVal 489
QY 1550 GAAGGTTTATTA 1561
D 490 GluTrpLeuLeu 493

RESULT 9
Q83RU5 SHIFL
ID Q83RU5_SHIFL PRELIMINARY; PRT; 493 AA.
AC Q83RU5_OTUCY8;
DT 01-JUN-2003 (TREMBLrel. 24, Created)
DT 13-SEP-2005 (TREMBLrel. 31, Last sequence update)
DE Putative synthase.
GN Name=yndC; OrderedLocusNames=S1116, SF1042;
OS Shigella flexneri.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Shigella.
OX NCBI_TaxID=623;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=301 / Serotype 2a;
RX MEDLINE=2272406; PubMed=12384590; DOI=10.1093/nar/gkf566;
RA Jin Q., Yuan Z., Xu J., Wang Y., Shen Y., Lu W., Wang J., Liu H.,
RA Yang J., Yang F., Zhang J., Zhang J., Yang G., Wu H., Qu D., Dong J.,
RA Sun L., Xue Y., Zhao A., Gao Y., Zhu J., Kan B., Ding K., Chen S.,
RA Cheng H., Yao Z., He B., Chen R., Ma D., Qiang B., Wen Y., Hou Y.,
RA Xu J.;
RT "Genome sequence of Shigella flexneri 2a: insights into pathogenicity
RT through comparison with genomes of Escherichia coli K12 and O157."
RL Nucleic Acids Res. 30:4432-4441(2002).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=2457T / ATCC 700930 / Serotype 2a;
RX MEDLINE=22590274; PubMed=12704152;
RX DOI=10.1128/IAI.71.5.2775-2786.2003;
RA Wei J., Goldberg M.B., Burland V., Venkatesan M.M., Deng W.,
RA Fournier G., Mayhew G.F., Plunkett G. III, Rose D.J., Darling A.,
RA Mau B., Perna N.T., Payne S.M., Runyen-Janecky L.J., Zhou S.,
RA Schwartz D.C., Blattner F.R.;
RT "Complete genome sequence and comparative genomics of Shigella
RT flexneri serotype 2a strain 2457T."
RL Infect. Immun. 71:2775-2786(2003).
DR EMBL; AE005674; RAN42664.2; -; Genomic DNA.
DR EMBL; AE016981; RAN16548.1; -; Genomic DNA.
DR GO; GO:0003824; F.catalytic activity; IEA.
DR GO; GO:0008152; P.metabolism; IEA.
DR InterPro; IPR001736; PLD.
DR Pfam; PF00614; PLDc; 2.
DR SMART; SM00155; PLDc; 2.
DR PROSITE; PS50035; PLD; 2.
KW Complete proteome.
SQ SEQUENCE 493 AA; 55912 MW; 495604AD934C6B92 CRC64;
```

Alignment Scores:

Pred. No.:	2,32e-53	Length:	493
Score:	1102.50	Matches:	221
Percent Similarity:	65.5%	Conservative:	83
Best Local Similarity:	47.6%	Mismatches:	143
Query Match:	38.7%	Indels:	17
DB:	2	Gaps:	5

US-10-665-990A-13 (1-1561) x Q83RU5_SHIFL (1-493)

```
QY 200 GGGCTATCCGATCTACTCTGCTCGACGACCCCGAGCCCTTGGCCCGCGCGCC 259
D 37 GlyGlnCysGlyLeuPheProLeuGluTyrSerLeuAlaPheAlaAlaArgTyrArg 56
QY 260 CTTATCGAATCTGCCGAAACACAGCTCGATTGCAATACATATTGGCGCAACACATT 319
D 57 LeuAlaGluMetAlaGluHisThrLeuAspValGlnTyrIleTrpGlnAspMet 76
```

```
QY 320 TCCGCGAGGCTGCTGTTCACACTCATGTACCTTCCCGAGAACCGCGGCTGGCGTACGC 379
D 77 SerGlyArgLeuLeuPheSerAlaLeuLeuAlaAlaLysArgGlyValArgValArg 96
QY 380 CTGCTGTTGGACGACAAACACACGCGGGTTCGACGATCTCTGCTCGCCCTCGACAGC 439
D 97 LeuLeuLeuAspAsnAsnThrProGlyLeuAspIleLeuArgLeuLeuAspSer 116
QY 440 CATCCCAATATCGAAGTGGCGCTGTTCACCCCTTCCTCGTCCGCAATATGGCGCCTC 499
D 117 HisProArgIleGluValArgLeuPheAsnProPheSerPheArgLeuLeuArgProLeu 136
QY 500 GGCTACCTGACCGATCTCCCGCTCAACCGCGCATGACACAAATCTTTACCGCC 559
D 137 GlyTyrIleThrAspPheSerArgLeuAsnArgMetHisAsnLysSerPheThrVal 156
QY 560 GACAAACCGCGCACCATCTCGCGGAGCGCAATATCGCGACGAATACTTCAAGTCGGT 619
D 157 AspGlyValValThrLeuValGlyGlyArgAsnIleGlyAspAlaTyrPheGlyAlaGly 176
QY 620 GAGGACACCGTTTCGCGACCTCGACATCTCGCCACCGCGAGCGTCTCGCGCAAGTA 679
D 177 GluGluProLeuPheSerAspLeuValMetAlaIleGlyProValValGluAspVal 196
QY 680 TCGCACGACTTCGACCGCTACTGGGCAAGCAATTCGCGCCCAACACCGCGCATATC 739
D 197 AlaAspAspPheAlaTyrTrpTyrCysLysSerValSerProLeuGlnValLeu 216
QY 740 CGCAGCGCAACATCGGCAAGGT-----CTTCAAGCACTCGGA 778
D 217 -----AspValProGluGlyGluMetAlaAspArgIleGluLeuProAlaSerTrp 233
QY 779 TACAACGACGAAACATCCAGACACGCGCTCTCGCTACCGCGAAACCGTCAACAGTCG 838
D 234 HisAsnAspAlaMetThrHis-----ArgTyrLeuArgLysMetGluSerSer 249
QY 839 CCCCTCTACCAAAATAACAGCGGAGCATCGACTGGGAGAGCGTCCAAACCGCGCTG 898
D 250 PropheIleAsnHisLeuValAspGlyThrLeuProLeuIleTrpAlaLysThrArgLeu 269
QY 899 ATCAGCAGACCCCTCGCAAAAGGACTCGACCGCAGCCCGCAACCCCGCATTCGCGG 958
D 270 LeuSerAspAspProAlaLysGlyGlyLysAlaLysArgHisSerLeuLeuProGln 289
QY 959 AGGCTGCAAGACCGCTCAAAACAGCCCAAGAGCGTCTATCTGTTTCAACCTATTTC 1018
D 290 ArgLeuPheAspIleMetGlySerProSerGluArgIleAspIleIleSerSerTyrPhe 309
QY 1019 GTCCCTACAAATCCGCGACAGACGCTGCAAACTGGTGCAGGCGGATACAGCTT 1078
D 310 ValProThrArgAlaGlyValAlaGlnLeuLeuArgMetValArgLysGlyValLysIle 329
QY 1079 ACCGTCTGACCAACTCGCTACAGCGACCGACGCTTCCCGCGCTCCATTCGCGCTACGTC 1138
D 330 AlaIleLeuThrAsnSerLeuAlaAlaAsnAspValAlaValHisAlaGlyTyrAla 349
QY 1139 AAATACCGAAACCGCTCTCAAGCGCGATCAAACTCTACGAGCTGCAACCCCAACCAT 1198
D 350 ArgTrpArgLysLysLeuLeuArgTyrGlyValGluLeuTyrGluLeuLysProThrArg 369
QY 1199 GCGTCCCGCCACCA--AAAGACAAAGCGCTGACCGCGAGCTCCGTAACCGACCTGCAT 1255
D 370 GluGlnSerSerThrLeuHisAspArgGlyIleThrGlyAsnSerGlyAlaSerLeuHis 389
QY 1256 GCCAAACCTTCATTGTGACGGCAAAACGCTATCTTCATCGGCTCATTTCAACCTCGACCC 1315
D 390 AlaLysThrPheSerIleAspGlyLysThrValPheIleGlySerPheAsnPheAspPro 409
QY 1316 GCTTCGCGACGCGCTCAATACCGAAATGGCGTGTGTCATCGAAAGCCCAAAATCCGAGNA 1375
D 410 ArgSerThrLeuLeuAsnThrGluMetGlyPheValIleGluSerGluThrLeuAlaGln 429
```

```

QY 1376 CAGATGGAGCGCACCCTCGCCGATACACACCCGGAATACGCTACCGGTTACCTCGAC 1435
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
430 LeuileAspLysArgPheileGlnSerGlnTyrAspAlaAlaTrpGlnLeuArgLeuasp 449
QY 1436 AAACAACACCGCTGCAATGGACGATCCGCGCCACCCGAAAA-----ACCTACCCGAAC 1489
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
450 ArgTrpGlyArgileAsnTrpValAspArgHisAlaLysLysGluValileLeuLysLys 469
QY 1490 GAAACCCGAAGCAAACTTTGGAACGCGATCGCGCGCAAAATCCTATCCCTGTCGCCCATC 1549
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
470 GluProAlaThrSerPheTrpLysArgValMetValArgLeuAlaSerileLeuProVal 489
QY 1550 GAAGTTTATTA 1561
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
490 GluTrpLeuLeu 493

RESULT 10
Q7AFAB_ECO57
ID Q7AFAB_ECO57 PRELIMINARY; PRT; 493 AA.
AC Q7AFAB;
DT 05-JUL-2004 (TREMBlrel. 27, Created)
DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)
DT 05-JUL-2004 (TREMBlrel. 27, Last annotation update)
DE Putative synthase.
GN OrderedLocusNames=Ecol1424;
OS Escherichia coli O157:H7.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=83334;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=O157:H7 / Sakai / RIMD 0509952 / EHEC;
RX MEDLINE=21156231; PubMed=11258796;
RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
RA Iida T., Takami H., Honda T., Sasaki C., Ogasawara N., Yasunaga T.,
RA Kuhara S., Shiba T., Hattori M., Shinagawa H.;
RT "Complete genome sequence of enterohemorrhagic Escherichia coli
RT O157:H7 and genomic comparison with a laboratory strain K-12.";
RL DNA Res. 8:11-22(2001)
DR EMBL; BA000007; BAB34847.1; -; Genomic DNA.
DR GO; GO:0003824; F:catalytic activity; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR001736; PLD.
DR Pfam; PF00614; PLDC; 2.
DR SMART; SM00155; PLDC; 2.
DR PROSITE; PS00035; PLD; 2.
SQ SEQUENCE 493 AA; 55928 MW; DEC407F83D7D1CFF CRC64;

Alignment Scores:
Pred. No.: 3e-53 Length: 493
Score: 1100.50 Matches: 220
Percent Similarity: 65.5% Conservative: 84
Best Local Similarity: 47.4% Mismatches: 143
Query Match: 38.6% Indels: 17
DB: 2 Gaps: 5

US-10-665-990a-13 (1-1561) x Q7AFAB_ECO57 (1-493)
QY 200 GGGTATCCGACATCTACTCTCGACGACCCCAAGCCCTTGGCGCCGCGCCGCC 259
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
37 GlyGlnCysGlyLeuPheProLeuGluLysSerLeuAspAlaPheAlaAlaArgTyrArg 56
QY 260 CTTATCGATCGCGACACAGCTCGATTGCAATCTACTACATTGGCGCAACGACATT 319
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
57 LeuAlaGluMetSerGluHisThrLeuAspValGlnTyrTrileTrpGlnAspMet 76
QY 320 TCCGCGAGCTGTGTTCAACTCATGTACCTTGGCGAGAACCGGGGTGCGGTACGC 379
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
77 SerGlyArgLeuLeuPheSerAlaLeuLeuAlaLysArgGlyValArgValArg 96
QY 380 CTGCTGTTGGACGACAAACACCGCGGGTGGACGATCTCTGCTCGCCCTCGACGC 439
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

```

Db 97 LeuLeuLeuAspAspAsnAsnThrProGlyLeuAspAspIleLeuArgLeuLeuAspSer 116
QY 440 CATCCCAATATCGAAGTCGCGCTGTTCAACCCCTTTCCTACGAAATGGCGGCATCTC 499
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
117 HisProArgIleGluValArgLeuPheAsnProPheSerPheArgLeuLeuArgProLeu 136
QY 500 GGCTACCTGACCGACTTCCCGCGCTCAACCGCGCATGTACACAAATCTTTTACCGCC 559
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
137 GlyTyrIleThrAspPheSerArgLeuAsnArgMetHisAsnLysSerPheThrVal 156
QY 560 GACAACCGCGCCACATATCTCGCGGACGCAATATCGCGACGAAATACTTCAAGTCGGT 619
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
157 AspGlyValValThrLeuValGlyGlyArgAsnIleGlyAspAlaTyrPheGlyAlaGly 176
QY 620 GAGGACACCGTTTTCGCGACCTCGACATCTCCGCCACCGCGAGCGTCTCGGCAAGTA 679
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
177 GluGluProLeuPheSerAspLeuAspValMetAlaIleGlyProValValGluAspVal 196
QY 680 TCGCACGACTTCGACCGCTACTGGGCAAGCCATTCCGCCCAACAACGCCACCGCATCATC 739
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
197 AlaAspAspPheAlaAaGlyTyrTrpTyrCysLysSerValSerProLeuGlnGlnValLeu 216
QY 740 CGCAGCGCAACATCGGCAAGGGT-----CTTCAAGCATCTCGGA 778
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
217 -----AspValProGluGlyGluMetAlaAspArgIleGluLeuProAlaSerTrp 233
QY 779 TACAACGACGAAACATCCAGACACGCGCTCTCGCGTACCGCGAAACCGTCAACAGTCG 838
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
234 HisAsnAspAlaMetThrHis-----ArgTyrLeuArgLysMetGluSerSer 249
QY 839 CCCCTCTACCAAAAAATACAGACGGGACGCGTACTGGCAGAGCGTCCAAACCGCGCTG 898
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
250 ProPheIleAsnHisLeuValAspGlyThrLeuProLeuIleTrpAlaLysThrArgLeu 269
QY 899 ATCAGCAGACCCCTGCAAAAGGACTCGACCGCGACCGCGCAACCCCGATTGCCGGG 958
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
270 LeuSerAspAspProAlaLysGlyGluGlyLysAlaLysArgHisSerLeuLeuProGln 289
QY 959 AGGCTGCAAGACGCGCTCAACAGCCCGCAAAAGCGTCTATCTGGTTTTCACCTATTTC 1018
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
290 ArgLeuPheAspIleMetGlySerProSerGluArgIleAspIleSerSerTyrPhe 309
QY 1019 GTCCCTACAAAAATCCGGCACAGACGCGTACTGCAAAACTGGTGAGGACGGCATAGACGTT 1078
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
310 ValProThrArgAlaGlyValAlaGlnLeuLeuArgMetValArgLysGlyValLysIle 329
QY 1079 ACCGTCCTGACCAACTCGCTACAGGCGACCGACGTTCCCGCGTCCATTCCGGTACGTC 1138
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
330 AlaIleLeuThrAsnSerLeuAlaAlaAsnAspValAlaValValHisAlaGlyTyrAla 349
QY 1139 AAATACCGAAACCGCTCTCAAGCCGCGCATCAAACTCTACGAGTGCACCAACCATC 1198
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
350 ArgTrpArgLysLysLeuLeuArgTyrGlyValGluLeuTyrGluLeuLysProThrArg 369
QY 1199 GCGCTCCCGCCCAACA---AAAGCAAAAGCGCTGACCGCGAGCTCCGTAACCCAGCTGCAT 1255
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
370 GluGlnSerSerThrLeuHisAspArgGlyIleThrGlyAsnSerGlyAlaSerLeuHis 389
QY 1256 GCCAAAACCTTCATTGTGGACCGCAACGCGATCTTCATCGGCTCATTTCAACCTGACCCC 1315
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
390 AlaLysThrPheSerIleAspGlyLysThrValPheIleGlySerPheAsnPheAspPro 409
QY 1316 CGTTCGCGCGCTCAATACGAAATGGCGTCTCATCCAAAGCCCAAAATCGCAGAA 1375
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
410 ArgSerThrLeuLeuAsnThrGluMetGlyPheValIleGluSerGluThrLeuAlaGln 429
QY 1376 CAGATGAGCGCACCTCGCGCATACCAACCCGGAATACCGCTACCGGTTACCTCGAC 1435
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
430 LeuileAspLysArgPheIleGlnSerGlnTyrAspAlaAlaTrpGlnLeuArgLeuasp 449
QY 1436 AAACAACACCGCTGCAATGGACGATCCGCGCCACCCGAAAA-----ACCTACCCGAAC 1489
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
450 ArgTrpGlyArgileAsnTrpValAspArgHisAlaLysLysGluIleValLeuLysLys 469

```

QY 1490 GAACCCGAGCCAACTTTGGAAACGATCGCGCAAAATCCTATCCTGCTGCCCATC 1549
 Db 470 GluProAlaThrSerPheTrpLysArgValMetValArgLeuAlaSerIleLeuProVal 489
 QY 1550 GAAGGTTTATTA 1561
 Db 490 GluTrpLeuLeu 493

RESULT 11
 Q57QL0 SALCH
 ID Q57QL0 SALCH PRELIMINARY; PRT; 528 AA.
 AC Q57QL0;
 DT 10-MAY-2005 (TrEMBLrel. 30, Created)
 DT 10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
 DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
 DE Putative phospholipase.
 GN Name=ymdC; OrderedLocustNames=SC1095;
 OS Salmonella cholerae-suis (Salmonella enterica).
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Salmonella.
 OX NCBI_TaxID=591;
 RN [1]
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
 RC STRAIN=SC-B67;
 RX PubMed=15781495;
 RA Chiu C.-H., Tang P., Chu C., Hu S., Bao Q., Yu J., Chou Y.-Y.,
 Wang H.-S., Lee Y.-S.;
 RT "The genome sequence of Salmonella enterica serovar Choleraesuis, a
 RT highly invasive and resistant zoonotic pathogen.";
 RL Nucleic Acids Res. 33:1690-1698(2005).
 DR EMBL; AE017220; AX65001.1; -; Genomic_DNA.
 KW Complete proteome.
 SQ SEQUENCE 528 AA; 60029 MW; D5175BE93D71CBDB CRC64;

Alignment Scores:

Score: 4.17e-53 Length: 528
 1098.00 Matches: 235
 Percent Similarity: 60.7% Conservative: 88
 Best Local Similarity: 44.2% Mismatches: 175
 Query Match: 38.5% Indels: 34
 DB: 2 Gaps: 8

US-10-665-990A-13 (1-1561) x Q57QL0_SALCH (1-528)

QY 35 CCGATGAAACACGACG-----CTCATTTCCCTTTATGCTCCTT 76
 Db 8 ProLeuLysSerProSerGlyArgPheGlnIleLeuLeuProValThrLeuCysLeuAsn 27
 QY 77 CTCTGTTTCATGTTCTTCATGGTTGCCCCCCTGGAGAACGACGGAAGCCGTCATTTC 136
 Db 28 ArgTyrThrLeuSerValMetMetLysLysLeuProGlyPheThrGlnAsp-----Tyr 45
 QY 137 AATACTTCCAAACCTGCTCCTCGGACAAACATCCTGCAATCCGGCACACC----- 187
 Db 46 LeuLeuSerLysAlaThrThrLeuProAspLysThrArgLeuGluArgAlaValGluPro 55
 QY 188 -----CCTCATACAAACGGGTATCCGACATCTACCTGCTCGACGACCCCGAACGCC 241
 Db 66 LeuCysAlaArgHisProGlyGluCysGlyIleLeuAlaLeuAspAsnSerLeuAspAla 85
 QY 242 CTTGCGCGCGCGCCCTTATCGAATCTGCGGAACACAGCCTCGATTGCAATACTAC 301
 Db 86 PheAlaAlaArgTyrArgLeuThrGluThrAlaAlaArgThrLeuAspValGlnTyrTyr 105
 QY 302 ATTTGGCGCAACGACATTTCCCGAGGCTGCTTCAACCTCATGTACCTGCTCGCGCAGAA 361
 Db 106 IleTrpGluAspMetSerGlyArgLeuLeuPheSerValLeuSerAlaAlaLys 125
 QY 362 CGCGCGTCCGGCTACGCTGCTGTGGACGACAAACACACGCGGGGTGGACCATCTC 421
 Db 126 ArgGlyValHisValArgLeuLeuAspAspAsnThrProGlyLeuAspThr 145

422 CTGCTCGCCCTCGACGACCATCCCAATATGAGTGGCGCTGTTCACCCCTTCGTCCTCA 481
 Db 146 LeuArgLeuLeuAspSerHisProAsnIleGluValArgLeuPheAsnProPheSerPhe 165
 QY 482 CGCAAAATCGCGCAGCTCGGCTAGCTACCGACACTTCCCGCCCTCAACCGCGCATGAC 541
 Db 166 ArgThrLeuArgAlaLeuGlyTyrLeuThrAspPheAlaArgLeuAsnArgMethHis 185
 QY 542 AACAAATCTTTACGCGGCAACCGCGCCACCATCTACTCGGCGGACGCAATATCGCGC 601
 Db 186 AsnLysSerTyrThrAlaAspGlyValValThrLeuValGlyArgAsnIleGlyAsp 205
 QY 602 GAATACTTCAAGTCGGTGAGGACACCGCTTTCCCGACCTGGACATCCTCGCCACCGC 661
 Db 206 AlaTyrPheGlyAlaGlyGluProLeuPheSerAspLeuAspValMetAlaIleGly 225
 QY 662 AGCGTCGTCGGCGAAGTATCGACGACTTCCACCGCTACTCGGCAAGCATTCGCGCC 721
 Db 226 ProValValAsnAspValAlaAsnAspPheGluArgTyrTrpArgCysSerSerValSer 245
 QY 722 AACGCCACGCGCATCATCGCGGCAACATCGCGAAGGTCTTCAAGCACTC----- 775
 Db 246 ThrLeuGlnGlnValLeuSerLeuSerGlu-----GlnGluLeuThrGln 260
 QY 776 -----GGATACAAACGACGAAACATCCAGACACGCGCTCCTCGCGC 814
 Db 261 ArgIleGluLeuProGluSerTrpTyrAsnAspGluIleThrArg-----Arg 276
 QY 815 TACCGCAAAACCGTCGACAGTCGCGCCCTCTACCAAAATAATACAGCGGACGATCGAC 874
 Db 277 TyrLeuHisLysLeuGluThrSerGlnPheMetAlaAspLeuAspArgGlyArgLeuPro 296
 QY 875 TGGCAGACGCTCAAAACCGCTGATCAGCAGACCCCTGCAAAAGAGTTCGCGCGCAC 934
 Db 297 LeuIleTrpAlaLysThrArgLeuLeuSerAspProSerLysGlyGluLysAla 316
 QY 935 CGCGCAAAACCGCGATTGCGGAGGCTCAAGACGCGCTCAAAACGCGCAAAAGC 994
 Db 317 GlnArgHisSerLeuLeuProGlnArgLeuPheAspValMetGlySerProThrGluArg 336
 QY 995 GTCTATCTGTTTCCCTATTTCTGCTCCTACAAATCCGGCACAGACGCTGCGCAAAA 1054
 Db 337 IleAspIleIleSerAlaTyrPheValProThrArgAlaGlyValAlaGlnLeuLeuAsn 356
 QY 1055 CTGTCGACGACGATAGACGTTACCTGCTCTGACCACTCGCTACAGCGGACCGACGTT 1114
 Db 357 LeuValArgLysGlyValLysIleAlaIleLeuThrAsnSerLeuAlaAlaAsnAspVal 376
 QY 1115 GCGCGCTCCATTCGCGCTAGCTCAATACCGAAACCGCTGCTCAAAACGCGCATCAAA 1174
 Db 377 AlaValValHisAlaGlyTyrAlaArgTrpArgLysLysLeuLeuArgTyrGlyValGlu 396
 QY 1175 CTCTACGAGCTGCAACCCCAACCATGCCC---GTCCCCGCCCAAAAGACAAAGCGCTGACC 1231
 Db 397 LeuTyrGluLeuLysProThrArgGluHisGluThrAlaValHisAspArgGlyLeuThr 416
 QY 1232 GGCAGCTCGTAACGAGCTGATGCCAAACCTTCATTGTGGACGGCAACGCACTCTTC 1291
 Db 417 GlyAsnSerGlySerSerLeuHisAlaLysThrPheSerIleAspGlySerLysValPhe 436
 QY 1292 ATCGGCTCATTCACCTCGACCCCGCTCCGACGCGCTCAATACCGAAATGGCGCTCGTC 1351
 Db 437 IleGlySerLeuAsnPheAspProArgSerThrLeuLeuAsnThrGluMetGlyPheVal 456
 QY 1352 ATCGAAACCCCAAAATCGCAGAACAGATGAGCGCACCCCTCGCGCATACCAACCCGAA 1411
 Db 457 IleGluSerGluThrLeuAlaThrLeuIleHisLysHisPheThrGlnSerGlnArgAsp 476
 QY 1412 TACGCTACCGGTTACCTCGACAAACACACCGCTGCAATGCGCATACCAACCCGAA 1465
 Db 477 AlaIleTrpGlnLeuArgLeuAspArgTrpGlyArgIleAsnTrpIleAspArgGlnGln 496
 QY 1466 GCCACCCGAAACACCTACCCGAAACCGAAACCGCAAACTTTTGAAACGCAATCGCCGA 1525


```

Db 497 GluGluGluLysValLeuLysGluProAlaThrArgPheTrpGlnArgValLeuVal 516
QY 1526 AAAATCCCTATCCCTCTCCCATCGAAGGTTTATTA 1561
Db 517 ArgLeuAlaAlaLeuProValGluTrpLeuLeu 528

RESULT 12
ID Q8X917 ECO57 PRELIMINARY; PRT; 493 AA.
AC Q8X917;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Putative synthase.
GN Name=yndC; OrderedLocusNames=z1680;
OS Escherichia coli O157:H7.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=83334;
RN [1]

NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=O157:H7 / EDL933 / ATCC 700927 / EHEC;
RX MEDLINE=21074935; PubMed=11206551; DOI=10.1038/35054089;
RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
RA Grotbeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamousis K.,
RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
RA Welch R.A., Blattner F.R.;
RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7.";
RL Nature 409:529-533(2001).
DR EMBL; AE005174; AAG55792.1; -, Genomic_DNA.
DR PIR; D85666; D85666.
DR GO; GO:0003824; P: catalytic activity; IEA.
DR GO; GO:0008152; P: metabolic process; IEA.
DR InterPro; IPR001736; PLD.
DR Pfam; PF00614; PLDC; 2.
DR SMART; SM00155; PLDC; 2.
DR PROSITE; PS50035; PLD; 2.
KW Complete proteome.
SQ SEQUENCE 493 AA; 55981 MW; DCC21D8AAB643457 CRC64;

Alignment Scores:
Pred. No.: 7.38e-53 Length: 493
Score: 1093.50 Matches: 219
Percent Similarity: 65.3% Conservative: 84
Best Local Similarity: 47.2% Mismatches: 144
Query Match: 38.3% Indels: 17
DB: 2 Gaps: 5

US-10-665-990a-13 (1-1561) x Q8X917_ECO57 (1-493)
QY 200 GGGGTATCCGACATCTACCTGCTCGACGACCCCGAAGCCCTTGCCGCGCGCGCC 259
Db 37 GlyGlnCysGlyLeuPheProLeuGluLysSerLeuAspAlaPheAlaAArgTyrArg 56
QY 260 CTTATCAATTCGCGAACAACAGCTCGATTGCAATCTACATTTGGCGGCAACAGCATT 319
Db 57 LeuAlaGluMetSerGluHisThrLeuAspValGlnTyrTrpGlnAspMet 76
QY 320 TCCGGCAGGCTGTGTTCAACCTCATCTACCTTCCCGCAGACGCGGTCGCGTACGC 379
Db 77 SerGlyArgLeuLeuPheSerAlaLeuLeuAlaAlaLysArgGlyValArgValArg 96
QY 380 CTGCTGTGGACGCAACAACACGCGGGTGGACGATCTCTCTGCTCGCCCTCGACGC 439
Db 97 LeuLeuLeuAspAsnThrProGlyLeuAspAspIleLeuArgLeuLeuAspSer 116
QY 440 CATCCCAATATCGAAGTCGCGCTGTTCAACCCCTCTGCTCTAGCAAAATCGCGGCACTC 499
Db 117 HisProAlaThrSerPheTrpLysArgValMetValArgLeuAlaSerIleLeuProVal 489

```



```

QY 1550 GNAGGTTTATTA 1561
Db 490 GluTrpLeuLeu 493

RESULT 13
Q9CKM5_PASMU PRELIMINARY; PRT; 514 AA.
ID Q9CKM5_PASMU PRELIMINARY; PRT; 514 AA.
AC Q9CKM5;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein PM1586.
GN OrderedLocustNames=PM1586;
OS Pasteurella multocida.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Pasteurella.
OX NCBI_TaxID=747;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Pm70; PubMed=11248100; DOI=10.1073/pnas.051634598;
RX MEDLINE=21145866;
RA May B.J., Zhang Q., Li L.L., Paustian M.L., Whittam T.S., Kapur V.;
RT "Complete genomic sequence of Pasteurella multocida Pm70.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:3460-3465 (2001).
DR EMBL; AE006196; AAK03670.1; -; Genomic DNA.
DR GO; GO:00046740; F:transferase activity; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR001736; PLD.
DR Pfam; PF00614; PLDC; 2.
DR SMART; SM00155; PLDC; 2.
DR PROSITE; PS50035; PLD; 2.
DR Complete proteome; Hypothetical protein.
SQ SEQUENCE 514 AA; 57882 MW; C65B0EB65D80F216 CRC64;

Alignment Scores:
Pred. No.: 8 43e-53 Length: 514
Score: 1092.50 Matches: 238
Percent Similarity: 61.7% Conservative: 78
Best Local Similarity: 46.5% Mismatches: 173
Query Match: 38.3% Indels: 23
DB: 2 Gaps: 8

US-10-665-990A-13 (1-1561) x Q9CKM5_PASMU (1-514)
QY 53 CTCATTCCCTTTATGCTCTCTCTGTTCAATGTTCTTANGTTGCCCCCACTGAA 112
Db 17 LeuPheSerLeuValIleLeuAlaLeuIleSerTyrGlnArgLeu-----ProThrSer 34
QY 113 GAACGGACGGAAGCGTCATTTCATTAATCTTCAACCTGCTCTCGGACACACATCTG 172
Db 35 AspArgProIleSerThrHis-----ProProLeuAsnProAsnGlyLeu 49
QY 173 CAATATCCCGCAC-----ACCCCTCATAACACAGCGCTATCCGACATCTAC 217
Db 50 LeuAlaArgHisIleLeuProGlnValAlaGlnHisProAsn---LeuThrGlyLeuTyr 68
QY 218 CTGCTCGACAGCCCCCAGAGCCCTTGGCGCGCGCGCGCCCTTATCGAATCTCCGAA 277
Db 69 ProLeuGlyAspGlyLysAlaPheLeuAlaArgLeuAlaLeuSerGluHisAlaGlu 88
QY 278 CACAGCTCGATTTCATACATACATTTGGCGCAACGACATTTCCGCGAGGCTGCTGTC 337
Db 89 HisThrLeuAspLeuGlnTyrTrpIleThrHisAsnAspValSerGlyHisLeuLeuLeu 108
QY 338 AACCTCATGTCCTTCCGCGAGAGCGCGTGTGCGGTACGCTGCTGTGTGGACACAAAC 397
Db 109 GlnSerLeuTyrLysAlaAlaValArgGlyValLysValArgLeuLeuLeuAspAsn 128
QY 398 AACACGCGGGTTGGAGATCTCTGCTGCGCCCTCGACGCCATCCCAATATCGAAGTG 457
Db 129 AsnThrLysGlyMetAspThrIleLeuAlaSerLeuLeuAlaHisProAsnIleGlnIle 148
QY 458 CGCCTGTTCAACCCCTTCGTCCTACGCAATGGCGCGCACTCGGCTACCTGACCGACTTC 517

```

```

Db 149 ArgLeuPheAsnProPheMetGlnArgGlnTyrArgTrpLeuGlyPheLeuSerAspPhe 168
QY 518 CCCCGCTCAACACCGCGCATGCACAAATTCCTTTACCGCGCACAAACCGCGCCACCAT 577
Db 169 PheArgLeuAsnArgMetHisAsnLysSerPheThrAlaAspGlyValMetSerIle 188
QY 578 CTGCGGAGCGCAATATCGCGGACGAATACTTTCAAAGTCGGTGGAGGACACCGTTTCCGC 637
Db 189 LeuGlyGlyArgAsnIleGlyAspGluTyrPheAspValGlyAsnGlyValLeuPheAla 208
QY 638 GACCTGGACATCTCGCCACCGGAGCGTCTCGCGAAGTATCGACGACTTCGACCGC 697
Db 209 AspLeuAspValAlaAlaThrGlyAlaValThrHisIleGlnThrAspPheAspArg 228
QY 698 TACTGGGCAACCCATTCGCCCCACACGCCGCGCATATCCGCGAGCGGCAACATCGCG 757
Db 229 TyrTrpAsnSerProSerSerTyrProLeuGluSerIleIleIleArgAspProIleThr 248
QY 758 AAGGCTCTTCAAGCACTCGGATACACGACGAAACATCCAGACGCGCTCTCGGCTAC 817
Db 249 ProPheAsnProLeuProAlaLeuAspAspGluThrGln-----ThrTyr 263
QY 818 CGCGAAACCGTCGAACAGTCGCCCTCTACCAAAAAATACAGCGGAGCGCATCGACATGG 877
Db 264 LeuLysGlnLeuThrGluLeuProPheAlaLysSerLeuLysAlaGlyThrLeuAlaPhe 283
QY 878 CAGACGCTCAAAACCGCGCTGATCAGCAGACACCCCTCGAAAGGACTCGACCGGACCGC 937
Db 284 ThrTrpAlaGluAlaGluLeuIleSerAspAspProLysLysAlaLeuGlyLysSerLeu 303
QY 938 CGAAACCGCGATTCGGGAGCGCTGCAAGCGGCTCAACAGCGCTCAACAGCCGAAAAAGCGTC 997
Db 304 IleGlnAspSerValLeuAlaHisIleAlaProThrMetLeuAsnAlaLysAsnLeu 323
QY 998 TATCTGGTTTCACCTATTTCCTCCCTACAAAAATCCCGGCACAGCGACTGGCAAACTG 1057
Db 324 IleIleValSerProTyrPheValProThrHisValGlyValAspPheLeuSerArgile 343
QY 1058 GTGACGAGCGCATAGACGTTACCGTCTGACCAACTCGCTACAGGCGACCGAGCTGCC 1117
Db 344 SerGlnThrGlyThrGlnValSerIleLeuThrAsnSerLeuGluAlaThrAspValSer 363
QY 1118 CCGCTCATTTCCGCTAGCTCAATACGAAACCGCTGCTCAAGCGCGGATCAAACTC 1177
Db 364 IleValHisSerGlyTyrAlaLysHisArgLysThrLeuLeuGlnLysGlnIleGlnLeu 383
QY 1178 TACGAGCTGCAACCAACCATCGCTGCCGCCCAAAAAAGCAAAAGGC-----CTGACC 1231
Db 384 TyrGluLeuLysPro---HisAlaThrIleGlnMetGluSerSerGlyHisLeuLeuLys 402
QY 1232 GGCAGCTCCGTAACAGCTCGCATGCCAAACCTTCATTGTGTGGCGGCAACGCAATCTTC 1291
Db 403 GlyAlaSerSerAlaSerLeuHisAlaLysThrPheThrLeuAspAsnArgTyrLeuPhe 422
QY 1292 ATCGCTCATTTCAACCTCGACCCCGTTCGCCAGGCTCAATACGAAATGGCGCTCGTC 1351
Db 423 ValGlySerPheAsnMetAspProArgSerAlaMetLeuAsnThrGluMetGlyLeuLeu 442
QY 1352 ATCGAAACCCCAAAATCGCAGAAACAGATGGAGCGCACCTCGCGATACACACCCGAA 1411
Db 443 IleAspSerProGluLeuAlaArgLeuLeuSerAspGlyLeuGlnGlnAsnGlnAlaAsn 462
QY 1412 TACGCTACCGGTTTACCTCGACAAACACACCGCTGCAATGGCAGATCCCGCCACC 1471
Db 463 TyrAlaPheSerValLysLeuAsnGluAlaGlnAlaLeuTyrTrpGluThrGlnGluAsn 482
QY 1472 CGAAAA-----ACCTACCCGACGAAACCCGAGCCAAACTTTGGAAGACGATCCCGCA 1525
Db 483 GlyLysTrpIleArgTyrGluAsnGluProHisThrSerTrpPheLysArgPheSerVal 502
QY 1526 AAAATCTCTATCCCTGCTGCCCATCGAAGTTTATTA 1561

```

Db 503 TrpCysLeuSerTrpLeuProValGluHisLeuLeu 514

RESULT 14

Q8ZQ28 SALTY

ID Q8ZQ28 SALTY PRELIMINARY; PRT: 494 AA.

AC Q8ZQ28;

DT 01-MAR-2002 (TrEMBLrel. 20, Created)

DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)

DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)

DE Putative phospholipase.

GN Name=ymdc; OrderedLocusNames=STM1148;

OS Salmonella typhimurium.

OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;

OC Enterobacteriaceae; Salmonella.

OX NCBI_TaxID=602;

RN [1]

RP NUCLEOTIDE SEQUENCE.

RC STRAIN=LT2;

RA MEDLINE=21534948; PubMed=11677609; DOI=10.1038/35101614;

RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,

RA Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,

RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,

RA Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,

RA Waterston R., Wilson R.K.;

RT "Complete genome sequence of Salmonella enterica serovar Typhimurium

RT LT2.";

RL Nature 413:852-856(2001).

DR EMBL; AE008749; AAL20078.1; -; Genomic DNA.

DR GO; GO:0016021; C:integral to membrane; IEA.

DR GO; GO:0016740; F:transferase activity; IEA.

DR GO; GO:0008152; P:metabolism; IEA.

DR GO; GO:0008654; P:phospholipid biosynthesis; IEA.

DR InterPro; IPR001736; PLD.

DR Pfam; PF00614; PLDC; 2.

DR SMART; SM00155; PLDC; 2.

DR PROSITE; PS00035; PLD; 2.

KW complete proteome.

SQ SEQUENCE 494 AA; 56217 MW; 571C01EFC4FCB288 CRC64;

Alignment Scores:

Pred. No.:	1.24e-52	Length:	494
Score:	1089.50	Matches:	221
Percent Similarity:	63.9%	Conservative:	77
Best Local Similarity:	47.4%	Mismatches:	147
Query Match:	38.2%	Indels:	21
DB:	2	Gaps:	5

US-10-665-990A-13 (1-1561) x Q8ZQ28 SALTY (1-494)

Qy 200 GGGCTATCCGACATCTACCTGCTCGACGACCCCGAAGCCCTTGGCGCCGCGCCG 259

Db 38 GlyGluCysGlyIleLeuAlaLeuAspAsnSerLeuAspAlaPheAlaAlaArgTyrArg 57

Qy 260 CTTATCGAATCTCGGACACAGCTCGATTGCAATCTACATTTGGCGCAACGACATT 319

Db 58 LeuThrGluMetAlaAlaArgLeuAspValGlnTyrIleTpgIuAspMet 77

Qy 320 TCCGGCAGGCTGCTTCAACCTCATCTACCTTCCGCGAAGCGGGTGGCGGTACGC 379

Db 78 SerGlyArgLeuLeuPheSerValLeuLeuSerAlaAlaLysArgGlyValHisValArg 97

Qy 380 CTGCTGTTGGACGACAAACACACGCGGGTGGACGATCTCTGCTCGCCCTCGACAGC 439

Db 98 LeuLeuLeuAspAsnAsnThrProGlyLeuAspAspThrLeuArgLeuLeuAspSer 117

Qy 440 CATCCCAATATCGAAGTGGCGCTTTCAACCCCTTCTGCTACGCAATATGGCGCACTC 499

Db 118 HisProAsnIleGluValArgLeuPheAsnProPheSerPheArgThrLeuAlaLeu 137


Qy 500 GGCTACTGACGACTTCCCGCGCTCAACCGCGGATGACACAAATCTTTACCGCC 559


Db 138 GlyTyrLeuThrAspPheAlaArgLeuAsnArgMetHisAsnLysSerTyrThrAla 157


Db 450 LeuAspArgTrpGlyArgIleAsnTrpIleAspArgGlnGlnGluGluValLeu 469
QY 1484 CCGAACGAAACCGAAGCCAAACTTTGGAAAGCGATCGCGCAAAATCTATCCTGCTG 1543
Db 470 LysLysGluProAlaThrArgPheTrpGlnArgValLeuValArgLeuAlaAlaLeu 489
QY 1544 CCCATCGAAGGTTTATTA 1561
Db 490 ProValGluTrpLeuLeu 495
RESULT 16
QSPG9 SALPA
ID QSPG9 SALPA PRELIMINARY; PRT; 495 AA.
AC QSPG9;
DT 01-FEB-2005 (TrEMBLrel. 29, Created)
DT 01-FEB-2005 (TrEMBLrel. 29, Last sequence update)
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE Hypothetical protein ymdC.
GN Name=ymdC; OrderedLocuNames=SPA1703;
OS Salmonella paratyphi-a.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
OX NCBI_TaxID=54388;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=ATCC 9150;
RX PubMed=1531882; DOI=10.1038/ng1470;
RA McClelland M., Sanderson K.E., Clifton S.W., Latreille P.,
RA Porwollik S., Sabo A., Meyer R., Bieri T., Ozersky P., McLellan M.,
RA Harkins C.R., Wang C., Nguyen C., Berghoff A., Elliott G.,
RA Kohlberg S., Strong C., Du F., Carter J., Kremizki C., Layman D.,
RA Leonard S., Sun H., Fulton L., Nash W., Miner T., Minx P.,
RA Delehaunty K., Fronick C., Magrini V., Nhan M., Warren W., Florea L.,
RA Spieth J., Wilson R.K.;
RT "Comparison of genome degradation in Paratyphi A and Typhi, human-
RT restricted serovars of Salmonella enterica that cause typhoid.";
RL Nat. Genet. 36:1268-1274 (2004).
DR EMBL; CP000026; AAU77627.1; -; Genomic DNA.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0016740; P:transferase activity; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR GO; GO:0008554; P:phospholipid biosynthesis; IEA.
DR InterPro; IPR001736; PLD.
DR Pfam; PF00614; PLDC; 2.
DR SMART; SM00155; PLDC; 2.
DR PROSITE; PS00035; PLD; 2.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 495 AA; 56310 MW; 8BAADE58874983A2 CRC64;
Alignment Scores:
Pred. No.: 1.24e-52 Length: 495
Score: 1089.50 Matches: 221
Percent Similarity: 63.9% Conservative: 77
Best Local Similarity: 47.4% Mismatches: 147
Query Match: 38.2% Indels: 21
DB: 2 Gaps: 5
US-10-665-990A-13 (1-1561) x QSPG9 SALPA (1-495)
QY 200 GGGCTATCCGACATCTACCTGCTCGACGACCCCAAGCCCTTCCGCCCGCGCGCC 259
Db 39 GlyLysGlyLeuAlaLeuAlaLeuAlaLeuAlaLeuAlaLeuAlaLeuAlaLeu 58
QY 260 CTTATCGAATCTCGGAACACAGCTCGATTTCGATTCATACATTTGGCCGACGACATT 319
Db 59 LeuThrGluMetAlaAlaArgThrLeuAlaValGlnTyrTrpIleTrpGluAspMet 78
QY 320 TCCGGCAGGCTGTGTTTCAACCTCATGTACCTTGGCCGAGAACCGCGCTGCGGTACGC 379
Db 79 SerGlyArgLeuLeuPheSerValLeuSerAlaAlaLysArgGlyValHisValArg 98
QY 380 CTGCTGTGGACGACAAACACGCGCGGTGGACGATCTCTGCTCGCCCTCGACAGC 439
|||||


Db 99 LeuLeuLeuAspAspAsnAsnThrProGlyLeuAspAspThrLeuA-gLeuLeuAspSer 118
QY 440 CATCCCAATATCGAAGTCGCGCTGTTTCAACCCCTTCTCTACGCAATATGCGCGCATC 499
Db 119 HisProAsnIleGluValArgLeuPheAsnProPheSerPheArgThrLeuArgAlaLeu 138
QY 500 GGGTACCTGACCGACTTCCCGCCCTCAACCGCGCATGCAACAAATCTCTTTACGCC 559
Db 139 GlyTyrLeuThrAspPheAlaA-gLeuAsnArgArgMetHisAsnLysSerTyrThrAla 158
QY 560 GACAAACCGCGCCACCTACTCGCGGACGCAATATCGCGACGAATCTTCAAAAGTCG 619
Db 159 AspGlyValValThrLeuValGlyArgAsnIleGlyAspAlaTyrPheGlyAlaGly 178
QY 620 GAGGACACCTTTTCGCGACCTGGACATCTCGCCACCGCGACGCTGCGGCAAGTA 679
Db 179 GluGluProLeuPheSerAspLeuAspValMetAlaIleGlyProValAlaAsnAspVal 198
QY 680 TCGCACGACTTCGACCGCTACTCGGCAAGCATTTCCGCCCAACGCGCGCATCATC 739
Db 199 AlaAsnAspPheGluArgTyrTrpArgCysSerSerValSerThrLeuGlnValLeu 218
QY 740 CGCAGCGGCAACATCGCAAGGTCTTCAAGCACTC----- 775
Db 219 SerLeuSerGlu-----GlnGluLeuThrGlnArgIleGluLeuProGlu 233
QY 776 ---GGATACACGACGAGAAACATCCAGACACGCGCTCTCGCTACGCGAAACGTCGAA 832
Db 234 SerTrpTyrAsnAspGluIleThrArg-----ArgTyrLeuHisLysLeuGlu 249
QY 833 CAGTCGCGCCCTTACCAAAAAATACAGACGGGACGATCGACTGGCAGAGCGTCCAAACC 892
Db 250 ThrSerGlnPheMetAlaAspLeuAspCysGlyThrLeuProIleLeuTrpAlaLysThr 269
QY 893 CGCTGATACGACACCCCTCGCAAAAGGACTCGACCGCGACCGCGCAACCGCGGATT 952
Db 270 ArgLeuLeuSerAspAspProSerLysGlyGluGlyAlaGlnArgHisSerLeuLeu 289
QY 953 GCCGGGAGGCTCAAGACGCGCTCAACACAGCCCGAAAAAGCGTCTATCTGTTTCAACC 1012
Db 290 ProGlnArgLeuPheAspValMetGlySerProThrGluArgIleAspIleSerAla 309
QY 1013 TATTTGCTCCTTACAAAAATCCCGCACAGCGCTCGCAAAACTGGTGACGAGCGGCATA 1072
Db 310 TyrPheValProThrArgAlaGlyValAlaGlnLeuLeuAsnLeuValArgLysGlyVal 329
QY 1073 GAGTTACGCTCTGACCACTCGCTACAGGCGACCGAGTTCGCGCGTCCATTCCGCG 1132
Db 330 LysIleAlaIleLeuThrAsnSerLeuAlaAlaAsnAspValAlaValHisAlaGly 349
QY 1133 TAGCTCAATACCGAAACCGTCTCAAGCGCGCATCAAACTCTACGAGTCCAAACC 1192
Db 350 TyrAlaArgTrpArgLysLysLeuLeuA-gTyrGlyValGluLeuTyrGluLeuLysPro 369
QY 1193 AACCATGCC---GTCCCGCCCAAAAGACAAAGGCTGACCGCGAGTCCGTAACCCAGC 1249
Db 370 ThrArgGluHisGluThrAlaValHisAspArgGlyLeuThrGlyAsnSerGlySerSer 389
QY 1250 CTGCATGCCAAACCTTCAATTGTGGACGGCAACGCATCTTCATCGCTCATTTCAACCTC 1309
Db 390 LeuHisAlaLysThrPheSerIleAspGlySerLysValPheIleGlySerLeuAsnPhe 409
QY 1310 GACCCCGTTCGCGACGCTCAATACGAAATGGCGTGTCTATCGAAGCCCAAAATC 1369
Db 410 AspProArgSerThrLeuLeuAsnThrGluMetGlyPheValIleGluSerGluThrLeu 429
QY 1370 GCAGAACAGATGAGCGCACCTCGCCGATACACACCCCAATACCGCTACCGCGTTACC 1429
Db 430 AlaThrLeuIleHisLysArgPheThrGlnSerGlnArgAspAlaAlaTrpGlnLeuArg 449
QY 1430 CTCGACAAACACAAACCGCTCGAATGGCACGAT-----CCCGCCACCGCAAAACCTAC 1483
Db 450 LeuAspArgTrpGlyArgIleAsnTrpIleAspArgGlnGlnGluGluLysValLeu 469
|||||


Qy 338 AACCTCATGTACCTTGCCGCGAGAACGCCGCGTGGCGGTACGCCCTGCTGTTGGACGACAAC 397





















```

QY 1121 GTCCATCCGGTACGTCAATACCGAAACCGTGTCTCAAAAGCGGCATCAAACTCTAC 1180
Db ||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
358 ValHisGlyTyrTyrAlaLysTyrArgGluGluLeuLeuGluSerGlyIleThrLeuTrp 377
QY 1181 GAGTGCACCAACCAATCGCTCCCGCCACAAAAGCAAAAGGCGTACCGGAGCTCC 1240
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
378 GluIleLysAlaSerAlaGluIle-----LysArgGlnTrpSerLeuIleGlySerGln 395
QY 1241 GTAAACGAGCTCATGCCAAAACCTTCATTTGTGACGGCAAAACGATCTTCATCGCTCA 1300
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
396 ArgAlaSerLeuHisAlaLysValIleValIleAspAlaAspPheIleValGlySer 415
QY 1301 TTCACCTCGACCCCGTCCCGACGGCTCAATACCGAAATGGGCGTGTATCGAAAGC 1360
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
416 MetAsnTrpAspProArgSerAlaTyrIleAsnThrGluMetAlaValHisIleThrHis 435
QY 1361 CCCAAATCGCAGACAGATGAGCGCACCTCGCCGATACACACCCGAAATACGCCCTAC 1420
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
436 ProGluTyrValLysLysAlaThrValGlnLeuAsnSerAlaLeuProIleAspAlaTyr 455
QY 1421 CGCGTTACCTCGCAACACACACCGCTGCAATGCAGATCCCGCCACCGCAAAACC 1480
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
456 TyrValGluIleArgGluGlyAsn---LeuHisTyrThrAspLeuGlyAsnGlyLysVal 474
QY 1481 TACCGAAGAACCGGACGCAAACTTTGGAAACCATCGCGCGCAAAATCTATCCCTG 1540
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
475 TyrAspSerGluProGluSerSerIleTrpArgSerIleGlyAlaTrpLeuSerGlyVal 494
QY 1541 CTGCCCATCGAGGTTTATTA 1561
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
495 LeuProIleGluGlyMetLeu 501

```

RESULT 21

```

Q8D861_VIBVU
ID Q8D861_VIBVU PRELIMINARY; PRT; 524 AA.
AC Q8D861
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Phospholipase D-family protein.
GN OrderedLocusNames=VVI3126;
OS Vibrio vulnificus.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Vibrio.
OX NCBI_TaxID=672;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=CMCP6;
RA Rhee J.H., Kim S.Y., Chung S.S., Kim J.J., Moon Y.H., Jeong H.,
RA Choy H.E.;
RL "Complete genome sequence of Vibrio vulnificus CMCP6."
RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE016807; AAO11447.1; -; Genomic DNA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR001736; PLD.
DR Pfam; PF00614; PLDC; 2.
DR SMART; SM00155; PLDC; 2.
DR PROSITE; PS00335; PLD; 2.
KW Complete proteome.
SQ SEQUENCE 524 AA; 59929 MW; 8C88CB38C8CADA7 CRC64;

```

Alignment Scores:

```

Pred. No.: 2.38e-40 Length: 524
Score: 869.50 Matches: 209
Percent Similarity: 56.5% Conservative: 88
Best Local Similarity: 39.7% Mismatches: 190
Query Match: 30.5% Indels: 39
DB: 2 Gaps: 14

```

US-10-665-990A-13 (1-1561) x Q8D861_VIBVU (1-524)

QY 29 ATATCCCGATGAAACA-----CCGAGCCTCATTTCCCTTTTATGCCTCTCTC 79

```

Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
17 LeuGlyProLeuArgThrPheIleHisLysAlaLeuThrTyrCysSerValMetLeuLeu 36
QY 80 TGTTCATGTTCTTCATGTTGGTCCCGCCCACTGAAACAAAGCGACGAAAGCGGTTCATTCAAT 139
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
37 CysAlaCysThrSer-----LeuGluAsnHisSerPro-----PheAsp 49
QY 140 ACTTCCAAACCTGCTCCTCTG-----GACAAATCCTCTGCAG-----ATCCGG 181
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
50 LysGlnProSerTyrGlnLeuGlyTyrGlnAlaAspSerArgLeuSerAlaTyrLeuAsn 69
QY 182 CACACCCCTCATAC---AACGGGCTATCCGACATCTACCTCTCGACGACCCCGACGAA 238
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
70 HisHisProGlnAspArgGluAsnLeuThrAlaPhePheProLeuAspLysGlyHisasp 89
QY 239 GCCCTTGGCGCCCGCGCCCTTATCGAATCTCCGAAACACAGCCTCGATTTGCAATAC 298
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
90 AlaLeuLeuAlaArgLeuAlaLeuIleGluAlaAlaAspLysThrLeuAspLeuGlnTyr 109
QY 299 TACATTTGGCCCAACGACATTTCCGGCAGGCTGTGTTCACCTCATGTACTCTGCCGCA 358
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
110 TyrIlePheArgAsnAspGluThrGlyGlnLeuLeuThrTyrArgLeuPheGluAlaAla 129
QY 359 GAACGCGCGTGGCGTACGCTGTGTGGACGACAAACACGCGGGGTGTGGACGAT 418
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
130 GlnArgGlyValArgValArgIleLeuLeuAspAspMetGlnLysArgAsn---AspGlu 148
QY 419 CTCTGTCTCGCCCTCGACGACCATCCCAATATCAAGTGGCGCTGTTCACCCCTTCGTC 478
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
149 GlyLeuAlaArgLeuSerAlaHisProAsnIleGlnIleArgLeuPheAsnProHisGln 168
QY 479 CTACGCAAAATGGCGGCGACTCGGCTACTACCGACCTTCCCGCGCTCAACGCGCGCATG 538
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
169 TyrArgThrAlaArgThrLeuAlaMetAlaSerAspPheSerArgLeuAsnArgArgMet 188
QY 539 CACAACAAATCCTTTACGCGCGCAACCGCCACCATCTACGCGGGACGCAATATCGGC 598
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
189 HisAsnLysSerLeuThrAlaAspSerValValAlaIleValGlyArgAsnIleGly 208
QY 599 GACGAATACTTCAAAGTGGTGGAGACACCGCTTTGGCGGACCTGGACATCTCTGCCACC 658
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
209 AsnGlnTyrPheSerValAsnSerProValGluPheGlyAspPheAspLeuMetLeuTyr 228
QY 659 GGCAGCGTCTCGCGGGAAGTATCGACGACTTCGACCGCTACTGGGCAAGCATTCCTGCC 718
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
229 GlyAsnSerValGluGlnThrAlaGluGlnPheAspLeuTyrTrpAsnSerLeuHisAla 248
QY 719 -----CACAACGCCGCGCATCATCCGCGCGGCAACATCGGC 757
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
249 ValProIleGluTrpLeuThrAspAsnProIleProValThrGlu----- 263
QY 758 AAGGTCTTCAAGCACTCGGATACAACGACGAAACATCC---AGACGCGCTCTCTGGCG 814
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
264 GluGluLeuGlnAla-----TrpLeuLysGluThrGlnLeuGluGlnLysPheThrGln 281
QY 815 TACCGCGAAACCGTCAACAGTCCGCCCTCTACCAAAAAAATACAGACGGGACGATCGAC 874
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
282 GlyArgTyrAspPheThrGlnLeuGlnLeuTyrGlnGlnPheThrAspLysSerLeuVal 301
QY 875 TGGCAGAGCGTCCAAACCCCGCTGATCAGCGACACCCCTGCAAAAAGGACTCGACCGCG 934
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
302 TrpTyrTrpGlyLysGlyGlnValTrpTyrAspLeuProAspLys-----ValAsp 318
QY 935 CGCGCAAAACCGCGATTTCCGGGAGGCTGCAAGACGCGCTCAACACAGCCGCAAAAAGC 994
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
319 ThrGlnAlaProGlnLeuAlaAspAsnLeuAlaSerLeuLeuArgThrValLysAspSer 338
QY 995 GTCTATCTGTTTACACCTATTTCTCCCTACAAATCCGCGACAGACGACTGCGCAAAA 1054
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
339 LeuValIleSerProTyrPheValProThrGluArgGlyThrGlnAlaLeuValGlu 358
QY 1055 CTGTGTCAGGACGGCATAGAGCTTACCCTCTGACCAACTCGTACAGCGGACCGACGTT 1114
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

```
Db 359 AlaAlaHisArgGlyValAspIleThrIleValThrAsnSerLeuAlaSerAsnAspVal 378
QY 1115 GCCGCGCTCCATTCGGCTACGTCAATACCGAAACCGCTCGCTCAAGCCGGCATCAA 1174
Db 379 PheAlaValHisGlyTrpTrpAlaLysTyrArgGlnAspLeuValGluAlaGlyLeuGln 398
QY 1175 CTCACGAGCTGCAACCAACCATCGCTCCCGCCGCAACAAAGACAAAGAGGCTGACCGGC 1234
Db 399 LeuTrpGluThrLysAlaSerAlaArgIleAspSer-----LysTrpSerPheThrGly 416
QY 1235 AGCTCCGTAACCGCTCGATGCCCAACCTTCATTGTGGAGCGGCAACGCGATCTTCATC 1294
Db 417 SerSerArgSerSerLeuHisAlaLysValLeuLeuIleAspHisArgLeuLeuPheAla 436
QY 1295 GGCTCATTCACCTCGACCCCGTTCGCGACGGCTCAATACCGAAATGGGGTGTGTATC 1354
Db 437 GlySerMetAsnTrpAspProArgSerAlaLeuLeuAsnThrGluMetAlaAlaValIle 456
QY 1355 GAAAGCCCCAAATCGCAGAAACAGATGGAGCGCACCTTCGCGGATACCCACCCGAATAC 1414
Db 457 GluHisProAspTyrValGlnSerSerGluAlaLysLeuProMetGlyLeuGluThrAsn 476
QY 1415 GCCTACCGCTTACCTCGCAACAAACACCGCTGCAATGGCAGCATCCGCGCCACCGGA 1474
Db 477 AlaTyrGlnValArgMet---LysAsnGlyGluValAlaTrpPheAspHisGlnSerGln 495
QY 1475 AAAACCTACCGCAACGAAACCGAACCAACTTTGGAAACGCATCGCGCAAAAATCCTA 1534
Db 496 ValTrpPheAspSerGluProGluAlaThrValTrpArgLysIleGlyAlaTrpPheAla 515
QY 1535 TCCTGTGTCGCCATCGAA 1552
Db 516 GlyIleLeuProIleGlu 521

RESULT 22
Q7MMB3_VIBVY
ID Q7MMB3_VIBVY PRELIMINARY; PRT; 554 AA.
AC Q7MMB3;
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein VV1159.
GN OrderedLocusNames=VV1159;
OS Vibrio vulnificus (strain VJ016).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Vibrio.
OX NCBI_TaxID=196600;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX PubMed=14656965; DOI=10.1101/gr.1295503;
RA Chen C.-Y., Wu K.-M., Chang Y.-C., Chang C.-H., Tsai H.-C.,
RA Liao T.-L., Liu Y.-M., Chen H.-J., Shen A.-B.-T., Li J.-C., Su T.-L.,
RA Shao C.-P., Lee C.-T., Hor L.-I., Tsai S.-F.;
RT "Comparative genome analysis of Vibrio vulnificus, a marine
RT pathogen."
RL Genome Res. 13:2577-2587(2003).
DR EMBL; BA00037; BAC93923.1; -; Genomic DNA.
DR GO; GO:0003824; F:catalytic activity; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR001736; PUD.
DR Pfam; PF00614; PLDC; 2.
DR PROSITE; PS00035; PUD; 2.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 554 AA; 63341 MW; 8A7980BB096AB6C8 CRC64;
```

```
Alignment Scores:
Pred. No.: 5.17e-40 Length: 554
Score: 863.50 Matches: 204
Percent Similarity: 55.3% Conservative: 90
Best Local Similarity: 38.3% Mismatches: 187
Query Match: 30.3% Indels: 51
DB: 2 Gaps: 12
```

```
US-10-665-990A-13 (1-1561) x Q7MMB3_VIBVY (1-554)
QY 29 ATATCCCGCATGAAAAACA-----CGAGCCTCATTTCCCTTTTATGCTCTCTTC 79
Db 47 LeuGlyProLeuArgThrPheIleHisLysAlaLeuThrTyrCysSerValMetLeuLeu 66
QY 80 TGTTCATGTTCTTTCATGG-----TTGCCCCCACTGGAAGAACGACGCGAA----- 124
Db 67 CysAlaCysThrSerPheGluAsnHisSerProPheAspLysGlnProSerTyrGlnLeu 86
QY 125 -----AGCGTCATTTCAATACTTCCAAACCTGCTCCTCTCG 160
Db 87 GlyTyrGlnAlaAspSerArgLeuSerAlaTyrLeuAsn----- 99
QY 161 GACAAACATCTGCAAAATCCGGCACACCCCTCATCAACACGGGCTATCCGACATCTACCTG 220
Db 100 -----HisHisProGlnAspArgGluLeuLeuThrAlaPhePhePro 113
QY 221 CTCGACGACCCCAAGAGCCCTTGGCCGCGCCGCTTATCGAATCTGCCGAACAC 280
Db 114 LeuAspLysGlyHisAspAlaLeuAlaArgLeuAlaLeuIleGluAlaAlaAspLys 133
QY 281 AGCTCATTTGCAATACTACTTTGGCGCAACACATTTCCGACAGCTGCTGTTCAC 340
Db 134 ThrLeuAspLeuGlnTyrTyrIlePheArgAsnAspGluThrGlyGlnLeuLeuThrTrp 153
QY 341 CTCATGTACCTTCCCGCAGAACCGGCGTGGCTAGCCCTGCTGTGTGGACGACAAAC 400
Db 154 ArgLeuPheGluAlaAlaGlnArgGlyValArgValArgIleLeuLeuAspAspMetGln 173
QY 401 ACGCGCGGTGGACGATCTCTGCTCGCTCGACAGCCATCCCAATATCGAAGTGGCG 460
Db 174 LysArgAsn---AspGluGlyLeuAlaArgLeuSerAlaHisProAsnIleGlnIleArg 192
QY 461 CTGTTCAACCCCTTCGCTACGCAATGGCGGCATCTCGCTACCTGACGACGATCCCC 520
Db 193 LeuPheAsnProHisGlnTyrArgThrAlaArgThrLeuAlaMetAlaSerAspPheSer 212
QY 521 CGCTCAACCGCGCATGCAACAAATCTTTACCGCGCCGCAACACCGGCGCCACCATCTC 580
Db 213 ArgLeuAsnArgArgMetHisAsnLysSerLeuThrAlaAspSerValAlaIleVal 232
QY 581 GGGCGGCGCAATATCGCGCAGCAATCTTCAAAGTGGTGAGGACACCGTTTCCGCGAC 640
Db 233 GlyGlyArgAsnIleGlyAsnGluTyrPheSerValAsnSerProValGluPheGlyAsp 252
QY 641 CTGACATCTCGCCACCGCGCAGCGTGTCTGGCGAAGATATCGCACGACTTCGACCGTAC 700
Db 253 PheAspLeuMetLeuTyrGlyAsnSerValGluGlnThrAlaGluGlnPheAspLeuTyr 272
QY 701 TGGGCAAGCCATTCGCGC-----CACACGCGCACCGCGCATCATC 739
Db 273 TrpAsnSerLeuHisAlaValProIleGluTrpLeuThrAspAsnProIleProValThr 292
QY 740 CGCAGCGGCAACATCGGCAAGGCTTTCAAAGCACTCGGATACACACGCAAAACATCC-- 796
Db 293 Glu-----GluGluLeuGlnAla-----TrpLeuLysGluThrGlnLeu 305
QY 797 AGACACGCGCTCTCGGCTACCGCGAAACCGTCGAACAGTCGCGCCCTCTACCAAAAATA 856
Db 306 GluGlnLysPheThrGlnGlyArgTyrAspPheThrGlnLeuGlnLeuTyrGlnGlnPhe 325
QY 857 CAGACGGGCGCATCGACTGGCAGAGCGTCCAAACCGCTGATCGACGACACCCCTGCA 916
Db 326 ThrAspLysSerLeuValTrpTyrTrpGlyLysGlyGlnValTrpTyrAspLeuProAsp 345
QY 917 AAAGGAGCTCGACCGCGACCGCCGCAAAACCGCTTCCCGGAGGCTCGCAAGACGCGCTC 976
Db 346 Lys-----ValAspThrGlnGluGlnGlnLeuAlaAspAsnLeuThrSerLeuLeu 362
QY 977 AAACAGCCCGGAAAAAGCGTCTATCTGTGTTACCTTATTCGTCCTACAAATCCGGC 1036
Db 363 ArgThrValLysAspSerLeuValLeuIleSerProTyrPheValProThrGluArgGly 382
```

Qy	1037	ACGAGCGCACTGGCAAAACCTGGTGCAGGACGGCATAGACGTTACCGCTCGACCAACTCG	1096
Db	383	ThrGlnAlaLeuValGluAlaAlaGlnArgGlyValGluIleThrIleValThrAsnSer	402
Qy	1097	CTACAGCGGACCGAGCTTGGCGCGCTCATTCGGCTACGTCAATACCGGAAACCGCTG	1156
Db	403	LeuAlaSerAsnAspValPheAlaValIleGlyTrpTyrAlaLysTyrArgGlnAspLeu	422
Qy	1157	CTAAAGCGCGCATCAAACTCTACGAGCTGCAACCCCAACCATGCGCTCCCGCCACAAAA	1216
Db	423	ValGluAlaGlyIleGlnLeuTrpGluThrLysAlaSerAlaArgIleAspSer-----	440
Qy	1217	GACAAAGCGCTACCGCGCAGCTCCGCTAACAGCGCTGCATGCCAAACCTTCATTGTGGAC	1276
Db	441	LysTrpSerPheThrGlySerSerArgSerSerLeuHisAlaLysValLeuLeuIleAsp	460
Qy	1277	GGCAAAACGCACTTTCATCGGCTCATCAACTCGACCCCGCTTCGCGACCGGTCAATACC	1336
Db	461	HisArgLeuLeuPheAlaGlySerMetAsnTrpAspProArgSerAlaLeuLeuAsnThr	480
Qy	1337	GAATGGGCTGTCATCGAAGCCCAAAATCCAGAAAGATGGAGCGCACCTCGCC	1396
Db	481	GluMetAlaAlaValIleGluHisProAspTyrValGlnSerSerGluAlaLysLeuPro	500
Qy	1397	GATACCACACCCGAATACGCTACCGGCTTACCCTCGACAAACACCAACCGCTGCAATGG	1456
Db	501	MetGlyLeuGluThrAsnAlaTyGlnValArgMet--LysAsnGlyGluValAlaTrp	519
Qy	1457	CACGATCCCGCACCCGAAACAACTACCCGAAACGACCGAAGCAACCTTTGGAAACGC	1516
Db	520	PheAspHisGlnSerGlnValTrpPheAspSerGluProGluAlaThrValTrpArgLys	539
Qy	1517	ATCGCCGCAAAAATCCTATCCCTGCTGCCCATCGAA	1552
Db	540	IleGlyAlaTrpPheAlaGlyIleLeuProIleGlu	551
RESULT 23			
Q8D800_VIBUO PRELIMINARY; PRT; 501 AA.			
AC	Q8D800;		
DT	01-MAR-2003 (TrEMBLrel. 23, Created)		
DT	01-MAR-2003 (TrEMBLrel. 23, Last sequence update)		
DT	01-MAR-2004 (TrEMBLrel. 26, Last annotation update)		
DE	Phospholipase D-family protein.		
GN	OrderedLocusNames=VV13197;		
OS	Vibrio vulnificus.		
OC	Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;		
OC	Vibrionaceae; Vibrio.		
OX	NCBI_TaxID=672;		
RN			
RP	NUCLEOTIDE SEQUENCE.		
RC	STRAIN=CMCP6;		
RA	Rhee J.H., Kim S.Y., Chung S.S., Kim J.J., Moon Y.H., Jeong H.,		
RA	Choy H.E.;		
RT	"Complete genome sequence of Vibrio vulnificus CMCP6.";		
RL	Submitted (DSC-2002) to the EMBL/GenBank/DBJ databases.		
DR	EMBL; AE016807; AA011511; -; Genomic DNA.		
DR	GO; GO:0016740; F:transferase activity; IEA.		
DR	GO; GO:0008152; P:metabolism; IEA.		
DR	InterPro; IPR001736; PLD.		
DR	Pfam; PF00614; PLDC; 2.		
DR	SMART; SM00155; PLDC; 2.		
DR	PROSITE; PSS0035; PLD; 2.		
KW	Complete proteome.		
SK	SEQUENCE 501 AA; 56571 MW; BF09ACE115F8D30B CRC64;		
Alignment Scores:			
Pred. No.:	5,82e-40	Length:	501
Score:	862.50	Matches:	197
Percent Similarity:	55.5%	Conservative:	91
Best Local Similarity:	38.0%	Mismatches:	198
Query Match:	30.2%	Indels:	33

[illegible]

	:::		:::	:::	:::	:::	
Db	326	ThrGluAlaGlyThrLeuGlyLeuLysArgPheValGluSerGlyValLysValThrIle	345				
Qy	1085	CTGACCAACTCGCTACAGGCACCGACGTTCGCCGGTCCATTCCGGCTACGTCAAATAC	1144				
Db	346	ValThrAsnSerLeuAlaSerAsnAspValPheAlaValHisGlyTrpTyrAlaLysTyr	365				
Qy	1145	CGAAACCGTGCTCAAAGCCGGCATCAAACTCTACGAGTGCACAACCAACCATCGCGTC	1204				
Db	366	ArgGluGluLeuLeuThrSerGlyIleThrLeuTrpGluIleLysAlaSerAlaGluIle	385				
Qy	1205	CCCCCACAAGAAGCAAGGCCTCACCGGAGCTCCGTAACCAGCTCGTAGCCAAAACC	1264				
Db	386	-----LysArgGlnTrpSerLeuIleGlySerGlnArgAlaSerLeuHisAlaLysVal	403				
Qy	1265	TTCATTGTGGAGCGCAACCGCATTTTCATCGGCTCATTCACCTCGACCCCGCTCCGCA	1324				
Db	404	IleValIleAspAlaAspPheIleValGlySerMetAsnTrpAspProArgSerAla	423				
Qy	1325	CGGCTCAATACCGAAATGGGGTTCGTATCGAAGAACCCCAAAATCGCAGAACAGATGGAG	1384				
Db	424	TyrIleAsnThrGluMetAlaValHisIleThrHisProGluTyrValLysGluAlaThr	443				
Qy	1385	CGCACCTTCGCGGATACCACACCCGAATACGCTTACCGCTTACCTTCGACAAACACAAC	1444				
Db	444	ValGlnLeuAsnSerAlaLeuProLysAspAlaTyrTyrValGluIleArgGluGlyAsn	463				
Qy	1445	CGCCTGCAATGGCAGCATCCCGCCACCCGAAAAACCTTACCGCAACCAACCCGAAGCCAAA	1504				
Db	464	--LeuHisTrpThrAspLeuGlyAsnGlyLysValTyrAspSerGluProGluSerSer	482				
Qy	1505	CTTTGGAAACGCATCGCCGGCAAAATCCTATCCCTGCTGCCCATAFGAGTTTATTA	1561				
Db	483	IleTrpGSerIleGlyAlaTrpLeuSerGlyValLeuProIleGluGlyMetLeu	501				

```

RESULT 24
Q4FTL4_9GAMM
ID Q4FTL4_9GAMM PRELIMINARY; PRT; 570 AA.
AC Q4FTL4;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Probable phospholipase D.
GN ORFNames=Psyc_0791;
OS Psychrobacter arcticum 273-4.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Moraxellaceae; Psychrobacter.
OX NCBI_TaxID=259536;
[1]
RN NUCLEOTIDE SEQUENCE.
RP STRAIN=273-4;
RC NASA Astrobiolgy Institute;
RG Ayala-del-Rio H.L., Chain P., Ponder M.A., Di Bartolo G., Ivanova N.,
RA Bergholz P.W., Hauser L., Land M., Bakermans C., Rodrigues D.,
RA Klappenschach J.A., Zarka D., Larimer F., Richardson P., Thomasow M.F.,
RA Tiedje J.M.;
RT "Complete sequence of Psychrobacter arcticum 273-4.";
RL Submitted (JUL-2005) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=273-4;
RA Copeland A., Lucas S., Lapidus A., Barry K., Dettler C., Glavina T.,
RC Hammon N., Israni S., Chain P., Di Bartolo G., Ivanova N., Hauser L.,
RA Land M., Larimer F., Pittluck S., Richardson P.;
RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; CP000082; AA218644.1; -; Genomic DNA.
SQ SEQUENCE 570 AA; 63466 MW; 0ECC7DAA9ECA3BC1 CRC64;

```

Alignment Scores:	
Pred. No.:	2,27e-39
Score:	852.00
Percent Similarity:	54.4%
Best Local Similarity:	36.2%
Length:	570
Matches:	196
Conservative:	99
Mismatches:	208

Query Match:	29.9%	Indels:	39
DB:	2	Gaps:	7
US-10-665-990A-13 (1-1561) x Q4FTL4_9GMM (1-570)			
Qy	52	CCTCATTTCCCTTTATGCTCTCTCTCTCTGTTTCATGTTCTTCATGGTTGCCCCCATCGGA	111
Db	30	ProLysGlnProHisLeuProGluSerGln-AlaLeuSerAlaA-rgValAsnAlaLeuTy	49
Qy	112	AGAACGGAGGAAAGCGTCATTTCAAATFACITTCAAACCTGTCCTCTGGAC-	163
Db	49	rGlnSerAspIleSerLysGlnAspAsnIleGluLysAspValGlnLysAspGlyAspSe	69
Qy	164	-----AACATCTCGAAATCGGCACACCCCTCATTAACAA	198
Db	69	rIleAspThrThrLysAspThrThrAsnValSerThrAspThrValThrProAsnThrAs	89
Qy	199	CGGCTATCCGACATCTACCTGCTCGACAGCCCCCAGAA	238
Db	89	pSerLeuAspLeuValAlaAlaIleSerGlnSerLysIleHisProAspLeuSerG	109
Qy	239	-----GCCCTTGCCGCCGCGCGCTTATCGAATC	270
Db	109	yTyrHisProIleValThrGlyAlaAsnAlaPheAlaSerArgSerIleLeuThrGlyMe	129
Qy	271	TGCGGAACACAGCCTCGATTTCGAATACTACATTTGGCGCAACGACATTTCCGGCAGGCT	330
Db	129	tAlaThrArgAsnIleAspAlaGlnTyTyIleTrpHisAspAspGlnAlaGlyGlnLe	149
Qy	331	GCTGTTCAAACCTCATGTACCTTCGCGAGAACGGCGTGGCGGTACGCTCTGTTGGA	390
Db	149	uMetLeuLysAspLeuTrpAspAlaAlaGluArgGlyValIleValArgLeuLeuAs	169
Qy	391	CGAC--AAACAACACGCGCGGTGGACGATCTCTGCTGCGCCCTCGACAGCATCCCAA	447
Db	169	pAspPheAsnAsnAsnAlaLysPheAspGlnHisLeuLeuArgPheAlaSerHisProAs	189
Qy	448	TATCGAAGTGGCTGTTCACCCCTTCGTCTCTACGCAAAATGGCGCGCACTCGGCTACCT	507
Db	189	nValAlaValArgIleIleAsnProLeuIleTyArgLysPheLysSerLeuAsnPheLe	209
Qy	508	GACGCACTTCCCGCGCTCAACCGCCGATCATCACAAATCTTTTACGCGCGACCAACCG	567
Db	209	uThrGlyLeuProArgIleAsnArgA-rgMeHisAsnLysSerMetIlePheAspGlyS	229
Qy	568	CGCCACCATCTCGGCGAGCGAATATCGGCGACGAATACTTTCAAAGTCGGTGAAGCAC	627
Db	229	nIleThrIleIleGlyArgAsnIleGlyAspGlyTyTrLeuSerAsnAspLysAsnSe	249
Qy	628	CGTTTTCCGCGACTGACATCTCGGCACCGGCGAGGCTGCTGGCGAAGTATCCGACGA	687
Db	249	rGlnPheAlaAspLeuAspValLeuLeuIleGlyLysValValAlaAspIleAspAsnSe	269
Qy	688	CTTCGACCGCTACTGGGCAAGCCATTTCGCCCCACACGCCACGCGCATCATC-	739
Db	269	rPheAlaSerTyTrpSerAlaProIleSerPheAspIleGlnThrLeuAlaThrLeuAs	289
Qy	740	-----CGCAGCGGCAACATCGGCAAGGCTTTCAGACACTCGGATACACAGCAAAAC	792
Db	289	pLysGlyGluThrThrAspPheValLysGlyLeuAspLysLeuLysThrAspGluLysSe	309
Qy	793	ATCAGACACGCGCTCTGCGC---TACCGGAAACCGTCGAAACAGTCCGCCCTTACCA	849
Db	309	rSerSerAsnGlySerLeuSerIleTyLysAlaAlaIleGluAspSerSerIleAspTh	329
Qy	850	AAAAATACAGCGGACCGCATTCGACTCGCAGAGCGTCCAAACCCGCTGATCAGCGCAC	909
Db	329	rAspLeuIleAsnLysArgValProPheArgTrpThrAspMetGlnPheLeuSerAspAs	349
Qy	910	CCCTGCAAAAGGACTCGACCGGACCGCGCAAAACCGCGGATTCGCGGAGGCTCAAGA	969
Db	349	pValGlyLeuLeuThrLysThrValProAlaAspThrAsnLeuValHisGlnLeuArgTh	369

```
QY 970 CCGCTCAACAGCCGAAAGACGCTCTATCTGCTTCCACCTATTTCGCTCCCTACAAA 1029
Db 369 rLeuLeuGlySerProSerLysLeuThrIleSerSerTyrPheValProThrLys 389
QY 1030 ATCCGGCAGACAGCCACTGGCAAAACCTGTGTGAGCAGCAGATAGACGTTACGCTCCTGAC 1089
Db 389 sAspGlyValAsnThrLeuAsnAlaLeuAlaGluSerGlyIleGluIleLysIleLeuTh 409
QY 1090 CAATTCGCTACAGCGACGAGCGTTCGCGCGCTCATTCGCGCTACGTCGAATACCGAAA 1149
Db 409 rAsnSerPheAspAlaThrAspValThrAlaValHisSerGlyTyrSerGlnTrpArgPr 429
QY 1150 ACCGCTGCTCAAAGCCGATCAAACTCTACGAGCTGCAACCAACCATCGCTCCCGC 1209
Db 429 sSerLeuLeuArgAlaGlyValLysIleTyrGluLeuLysSerThrAlaSerGluGluLy 449
QY 1210 CACAAAAGACAAAGCCCTGACCGGCG-----AGCTCCGTAAACCGCTGCGATGCCAAAAC 1263
Db 449 sArgGluAsnLysLeuTrpLysGlyArgSerGlnSerSerThrSerLeuHisAlaLysAl 469
QY 1264 CTTCAATGTGGACGCAACAGCATCTTCATCGGCTCATTCACCTGACCGCCCGCTCCGCG 1323
Db 469 aPheAlaValAspAspTyrGlnValPheIleGlySerTyrAsnValAspProArgSerAl 489
QY 1324 ACGGCTCATACCGAAATGGCGCTGCTCATCGAAAGCCCAAAATCGCAGACAGATGGA 1383
Db 489 aAsnIleAsnThrGluMetGlyValIleIleAsnAspAspGluLeuAlaThrGlnLeuHi 509
QY 1384 GCGCACCTCGCGCATACACACCCGGAATACCGCTACCGCGTTACCTTCGACAAAACACAA 1443
Db 509 sAspAlaLeuSerAspAspLeuLeuAsnGlnAlaTyrGluValLysLeuLeuGluAsnGl 529
QY 1444 CCGCTGCAATGGCAC-----GATCCGCGCCACCCGAAACCTACCGAAGAACCCGA 1497
Db 529 yAsnLeuGlnTrpHisThrMetGluLysAspLysLysValIleTyrAspSerGluProAr 549
QY 1498 AGCCAACTTTGGAACGCATCGCGCAAAATCTATCCCTGCTGCCCATCAGAGTTT 1557
Db 549 gValAlaValSerAspHisValTrpLeuThrIleMetSerTrpLeuProIleAspTrpLe 569
QY 1558 ATTA 1561
Db 569 uLeu 570

RESULT 25
QBIMP4_XANAC
ID QBIMP4_XANAC PRELIMINARY; PRT; 520 AA.
AC QBIMP4;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-WAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Cardiolipin synthase.
GN Name=cls;
OS Xanthomonas axonopodis (pv. citri).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC Xanthomonadaceae; Xanthomonas.
OX NCBI_TaxID=92829;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=306 / ATCC 13902 / XV 101;
RX MEDLINE=22022145; PubMed=12024217; DOI=10.1038/417459a;
RA da Silva A.C.R., Ferro J.A., Rainach F.C., Farah C.S., Furlan L.R.,
RA Quaggio R.B., Monteiro-Vitorello C.B., Van Sluys M.A.,
RA Almeida N.F. Jr., Alves L.M.C., do Amaral A.M., Bertolini M.C.,
RA Camargo L.E.A., Camarotte G., Cannavan F., Cardozo J., Chambergo F.,
RA Ciapina L.P., Cicarelli R.M.B., Coutinho L.L., Cursino-Santos J.R.,
RA El-Dorcy H., Faria J.B., Ferreira A.J.S., Ferreira R.C.C., Gruber A.,
RA Ferro M.I.T., Formighieri E.F., Franco M.C., Greggio C.C., Gruber A.,
RA Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.W., Lemos M.V.F.,
RA Locati E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,
RA Martins E.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,
RA Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,
RA Pereira H.A., Rossi A., Sena J.A.D., Silva C., de Souza R.F.,
```

```
RA Spinola L.A.F., Takita M.A., Tamura R.E., Teixeira E.C., Tezza R.I.D.,
RA Trindade dos Santos M., Truffi D., Tsai S.M., White P.F.,
RA Setubal J.C., Kitajima J.P.;
RT "Comparison of the genomes of two Xanthomonas pathogens with differing
RT host specificities.";
RL Nature 417:459-463(2002).
DR EMBL; AE011929; AAM37716.1; -; Genomic DNA.
DR GO; GO:0003844; F: catalytic activity; IEA.
DR GO; GO:0008152; P: metabolism; IEA.
DR InterPro; IPR001736; PLD.
DR Pfam; PF00614; PLDC; 2.
DR SMART; SM00155; PLDC; 2.
DR PROSITE; PS00035; PLD; 2.
KW Complete proteome.
SQ SEQUENCE 520 AA; 57436 MW; 61A2DE5B0349A2FA CRC64;

Alignment Scores:
Pred. No.: 2,73e-39 Length: 520
Score: 850.50 Matches: 206
Percent Similarity: 52.7% Conservative: 80
Best Local Similarity: 37.9% Mismatches: 182
Query Match: 29.8% Indels: 75
DB: 2 Gaps: 15

US-10-665-990A-13 (1-1561) x QBIMP4_XANAC (1-520)
QY 56 ATTTCCTTTATGCTCTCTCTCTCTGTTTCATGT----- 88
Db 12 IleAlaLeuValLeuValLeuAlaSerAlaLeuSerLeuTyrGlyTyrGlyArgPhe 31
QY 89 -----TCTTCATGTTGCCCTGCAAGAACGCGGAA 124
Db 32 AlaAspArgGlnArgGlyProValSerHisValLeuProAla----- 45
QY 125 AGCGCTCATTTCAATACCTTCCAAACCTCTCTCTCGGACAAACATC-----CTGCAA 175
Db 46 -----ThrAlaValAlaThrProIleAspLysValAlaProLeuGln 60
QY 176 ATCGGCGACACCCCTCATACACACGGGTATCCGACATCTACCTCTCGACGACCCCGAC 235
Db 61 GlnAlaHisAlaAspGln-----ThrGlyMetValIleLeuProAspAsnVal 76
QY 236 GAAGCCCTTGGCGCGCGCGCTTATCGAATCTGCGGAAACACAGCCTCGATTTGCAA 295
Db 77 AspAlaPheAlaValArgAlaLeuThrAlaArgAlaAlaGlyArgSerLeuAspLeuGln 96
QY 296 TACTACATTTGGCGCAACGACATTTCCGCGAGGTGCTGTTCACCTCATGTACCTTGC 355
Db 97 TyrTyrIleTrpHisAlaAspPheThrGlyAsnLeuLeuHisAsnGluLeuArgAla 116
QY 356 GCAGAACCGCGGTGGGTACGCTGCTGTGGACGACACACACACGCGCGGTGGAC 415
Db 117 AlaAspArgGlyValArgValArgLeuLeuLeuAspMetAsnIleHisGlySerAsp 136
QY 416 GATCTCTCTGCTCGCCCTCGACAGCCATCCCAATATCGAAGTGCCTGTTCACCCCTTC 475
Db 137 SerValLeuAlaAlaLeuAspSerHisProLeuIleGluIleArgLeuPheAsnProThr 156
QY 476 GTCCTACGCAAA-----TGGCGCGCATCTCGGCTACCTGACCGCATCTCCCGCGCTC 526
Db 157 ArgAlaArgGluGlyThrLeuMetArgGlyValGluMetValLeuArgMetPheSerIle 176
QY 527 AACCGCCGATGACACAAATCCTTTACCGCGCAACACCGCGCACCATCTACTCCGCGGA 586
Db 177 AsnArgArgMetHisAsnLysAlaTrpIleAlaAspGlyArgIleAlaValValGlyGly 196
QY 587 CGCAATATCGCGCGCAATATCTTCAAAGTCGCTGAGGACACCGTTTTTCCCGCACTTGGAC 646
Db 197 ArgAsnValGlyAspGluTyrPheAspAlaAlaArgAspThrAsnPheMetAspMetAsp 216
QY 647 ATCTCTCGCCACCGCGAGCGTCTCGCGGAAAGTATCGCAGCATCTCGACGCTTCTGGGCA 706
Db 217 AlaAlaLeuMetGlyProAlaValGlyGlnAlaGluGlnValPheAspAlaTyrTrpAsn 236
```



```
Db 199 AsnIleGlyAspGluTyrPheGlyAlaSerGluGluMetAsnPheGlyAspPheAspLeu 218
Qy 650 CTCGCCACCGCAGCGTCGTGGCGAAGTATCGACGACTTCGACCGCTACTGTGGCGAAGC 709
Db 219 TrpAlaIleGlyProValValGlnGluLeuSerArgHisPheAspThrTyrTrpAsnSer 238
Qy 710 CATTCCGCCCAACAGCCACCGGCATCATCCGCGAGCGGCAACATCGGCGAAGGGTCTTCAA 769
Db 239 GluIleAlaTyrProIleSerValLeuValLys-----GlyPheLysProThrLeuGlu 256
Qy 770 GCACCTCGGATACACAGCGAAGACATCCAGACACGCGCTC-----CTGCGC 814
Db 257 AspLeuGlnLysLeuLysAspAlaAlaAlaIleThrGluAlaGluLysThrGln 276
Qy 815 TACCGCGAAACCGTCGAACAGTCGCGCTCTACCAAAAATA-----CAGACGGGAGCG 868
Db 277 TyrAlaAspAlaLeuLysGluThrProIleValLysLysPheThrHisGluProMetLys 296
Qy 869 ATCGACTGGCAGAGCGTCCAAACCGCCTGATCAGCGACACCCCTGCAAAA---GGACTC 925
Db 297 LeuTyrTrpGlyLysAlaAsp-----ValValMetAspProProGluLysPheHisGln 314
Qy 926 GACCGCGACCGCGCAACCGCGATTCGCGGAGGCTGCAAGAGCGCTCAACACAGGCC 985
Db 315 AspSerLysAspGlnValAspAsnLeuAlaArgGlnLeuTyrProLeuIleGluLysThr 334
Qy 986 GAAAAAGCGTCTATCTGTTTTCACCCCTATTTCGTCCTACAAAAATCCGGCACACGCA 1045
Db 335 GluLysGluLeuIleLeuValSerProTyrPheValProGlyLysLysGlyValLysPhe 354
Qy 1046 CTGGCAAACTGGTGCGAGGAGCGATGACGTTACCTCCCTGACCAACTCCTCAGCGC 1105
Db 355 PheLysHisLeuAsnAspArgGlyValGlnSerLeuValLeuThrAsnSerLeuAlaSer 374
Qy 1106 ACCGAGCTGCGCGCTCATTCGCGCTACGTCGCAATACCGAAACCCGCTGCTCAAAGCC 1165
Db 375 SerAspValAlaThrValPheSerGlyTyrLysGlyTyrArgLysAspLeuLeuAspGln 394
Qy 1166 GGCATCAAACTCTACGAGCTCAACCCCAACCATCGCTCCCGCCGACAAAGAGCAAAAGC 1225
Db 395 GlyValSerLeuTyrGluLeuLysProAsn---SerProThrThrThrProLysLysAsn 413
Qy 1226 CTGACCGCGACGTCGTGAACAGC-----CTGCATGCAAAACCTTCATTGTGGACGCG 1279
Db 414 ArgValGlySerSerPheSerSerAlaGlyLeuHisGlyLysIlePheValPheAspArg 433
Qy 1280 AAACGCATCTTCATCGGCTCATTCACCTGACCCCGCTTCGCGACGCTCAATACCGAA 1339
Db 434 LysLysValPheValGlySerMetAsnLeuAspProArgSerAlaThrLeuAsnSerGlu 453
Qy 1340 ATGGCGCTGTCATCGAAAGCCCCAAAATCGCAGAACATGAGCGCACCCCTCGCGAT 1399
Db 454 MetGlyValValValAspSerProGluLeuAlaGluMetIleSerThrAsnLeuIleAla 473
Qy 1400 ACCACACCGGATACGCTACCGGTTTACCTTCGCAACAAACACACCGCTGCAATGG--- 1456
Db 474 HisLeuArgArgAspSerTyrGlnValLeuLeuAsnGluLysAsnAsnLeuIleTrpLys 493
Qy 1457 -----CAGCATCCCGCCACCGGAAAAACCTACCGCAACCGAACCGGCAAACTTTGG 1510
Db 494 ThrThrAspAspCysGlyLeuGluHisValPheSerLysAspProGluThrSerTrpTrp 513
Qy 1511 AAACGATCGCGCAAAAATCTATCTCTGCTGCGCCATCGAAGGTTTATTA 1561
Db 514 LysArgPheLysAlaSerLeuSerAlaIlePheIleProGluSerTrpLeu 530

RESULT 27
Q5QUY8_IDILO
ID Q5QUY8_IDILO PRELIMINARY; PRT; 464 AA.
AC Q5QUY8;
DT 01-FEB-2005 (TReMBLrel. 29, Created)
DT 01-FEB-2005 (TReMBLrel. 29, Last sequence update)
```

```
DT 01-FEB-2005 (TReMBLrel. 29, Last annotation update)
DE Phospholipase D family protein.
GN OrderedLocusNames=Ilu144;
OS Idiomarina loihiensis.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Alteromonadales;
OC Idiomarinaceae; Idiomarina.
OX NCBI_TaxID=135577;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=L2-TR / DSM 15497 / ATCC BAA-735;
RX PubMed=15596722; DOI=10.1073/pnas.0407638102;
RA Hou S., Saw J.H., Lee K.S., Freitas T.A., Belisle C., Kawarabayasi Y.,
RA Donachie S.P., Pika A., Galperin M.Y., Koonin E.V., Makarova K.S.,
RA Omelchenko M.V., Sorokin A., Wolf Y.I., Li Q.X., Keum Y.S.,
RA Campbell S., Denery J., Aizawa S.-I., Shibata S., Malahoff A.,
RA Alam M.;
RT "Genome sequence of the deep-sea gamma-proteobacterium Idiomarina
RT loihiensis reveals amino acid fermentation as a source of carbon and
RT energy.";
RL Proc. Natl. Acad. Sci. U.S.A. 101:18036-18041(2004).
DR EMBL; AE017340; AA082284.1; -; Genomic DNA.
DR GO; GO:0003824; F:catalytic activity; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR01736; PLD.
DR Pfam; PF00614; PLDC; 2.
DR SMART; SM00155; PLDC; 2.
DR PROSITE; PS00335; PLD; 2.
KW Complete proteome.
SQ SEQUENCE 464 AA; 53557 MW; 59A8BFA13DEB45F7 CRC64;

Alignment Scores:
Pred. No.: 4,81e-39 Length: 464
Score: 846.00 Matches: 186
Percent Similarity: 56.6% Conservative: 76
Best Local Similarity: 40.2% Mismatches: 173
Query Match: 29.7% Indels: 28
DB: 6

US-10-665-990A-13 (1-1561) x Q5QUY8_IDILO (1-464)
Qy 218 CTGCTCGACGACCCCGCAGAACCCCTTTCGCCCGCCGCGCTTATCGAATCTGCCGA 277
Db 15 LeuLeuGluAsnAlaHisArgAlaLeuIleGluArgValArgLeuLeuArgGlnAlaArg 34
Qy 278 CACAGCTCGATTTCGCAATACTACTATTGGCGCAACGACATTCGGCAGCGTGTCTTC 337
Db 35 HisHisIleAlaLeuGlnTyrTyrLeuTrpArgProAspSerSerGlyLeuThrLeu 54
Qy 338 AACCTCATGTACCTTCGCGAGAACCGCGCTGCGCTGCTGCTGCGACGACAAC 397
Db 55 LysGluLeuLeuAspAlaValGluArgGlyValArgValAspLeuLeuAspLeu 74
Qy 398 AACACGCGCGGGTTGGACGATCTCTGCTGCCCTCGACAGCCATCCCAATATCGAAGTG 457
Db 75 HisSerLysProLeuGluProLeuLeuArgAspLeuSerSerArgAlaAsnPheAsnIle 94
Qy 458 CGCTGTTTCAACCCCTTCGTCCTACGCAATTCGGCGGCGACTCGGCTACTGACCGACTTC 517
Db 95 ArgLeuPheAsnProPheLysHisArgArgTrpArgTrpLeuAsnTrpLeuThrAspPhe 114
Qy 518 CCCGCTCAACCGCGCATGCACAAATCTTTTACCGCGCACACCGCGCCACCATA 577
Db 115 LysArgMetAsnArgArgMetHisAsnLysSerMetLeuValAspAlaGluValGlyIle 134
Qy 578 CTCGGCGGACGCAATATCGCGGCGAATATCTTCAAAAGTCGGTGAGGACACCGTTTTTCGC 637
Db 135 ValGlyGlyArgAsnValGlyAsnGluTyrPheGlyThrHisAlaGlyGlnLeuPheSer 154
Qy 638 GACCTGGACATCTCGCCACCGCGAGCTGCTCGCGGAAGTATCGACGACTTCGACCGC 697
Db 155 AspLeuAspValIleAlaLysGlyLysIleValLysThrLeuLeuAspTrpArgGln 174
Qy 698 TACTGGGCAAGCCATTTCGCCCCACAAACGCCACGCGCATCATCCGCGAGCAACATCGGC 757
```



```
Qy 854 ATACAGAGGGACGC-----ATCGACTGGCAGAGCGTCCAAACCGCGCTGATCAGCGAC 907
Db 854 ATACAGAGGGACGC-----ATCGACTGGCAGAGCGTCCAAACCGCGCTGATCAGCGAC 907
Qy 286 LeuMetGlnGlyAsnArgProValHisTrpValGln---GlnAlaArgIleValAlaAsp 304
Db 286 LeuMetGlnGlyAsnArgProValHisTrpValGln---GlnAlaArgIleValAlaAsp 304
Qy 908 ACCCTCGCAAAAGG-----CTCGACCGCAGCGCGCAAAACCGCGCTGATCAGCGAC 955
Db 908 ACCCTCGCAAAAGG-----CTCGACCGCAGCGCGCAAAACCGCGCTGATCAGCGAC 955
Qy 305 ProProGlnGlyAlaGluGlyAlaProProGlnProAspTrpMetThrProIleLeuIle 324
Db 305 ProProGlnGlyAlaGluGlyAlaProProGlnProAspTrpMetThrProIleLeuIle 324
Qy 956 GGGAGGCTGCAAGACGGCTCAAAACAGCCGCAAAAGCGTCTATCTGTTTACCTCAT 1015
Db 956 GGGAGGCTGCAAGACGGCTCAAAACAGCCGCAAAAGCGTCTATCTGTTTACCTCAT 1015
Qy 325 GlyGluMetAlaAsnAla-----ArgArgGluLeuLeuLeuLeuLeuSerProTyr 340
Db 325 GlyGluMetAlaAsnAla-----ArgArgGluLeuLeuLeuLeuLeuSerProTyr 340
Qy 1016 TTCCTGCTCAAAATCGGCACAGACGCTGCAAAACCTGGTGCAGGCGGATAGAC 1075
Db 1016 TTCCTGCTCAAAATCGGCACAGACGCTGCAAAACCTGGTGCAGGCGGATAGAC 1075
Qy 341 PheValProGlyAspGlyMetArgTrpIleGlyLeuArgGlyAspValArg 360
Db 341 PheValProGlyAspGlyMetArgTrpIleGlyLeuArgGlyAspValArg 360
Qy 1076 GTTACCGTCTGACCAACTCGCTCAGCGCAGCGCTGCGCGCTCCATTCGGGTAC 1135
Db 1076 GTTACCGTCTGACCAACTCGCTCAGCGCAGCGCTGCGCGCTCCATTCGGGTAC 1135
Qy 361 ValSerValLeuThrAsnSerLeuAlaAlaAsnAspValValAlaValHisSerGlyTyr 380
Db 361 ValSerValLeuThrAsnSerLeuAlaAlaAsnAspValValAlaValHisSerGlyTyr 380
Qy 1136 GTCAAAATACCGAAACCGCTCTCAAAACCGCGATCAAACTCTACGAGCTGCAACCCAAAC 1195
Db 1136 GTCAAAATACCGAAACCGCTCTCAAAACCGCGATCAAACTCTACGAGCTGCAACCCAAAC 1195
Qy 381 AlaAspTyrArgIleProLeuLeuGlnGlyValArgLeuHisGluLeuLeuPro--- 399
Db 381 AlaAspTyrArgIleProLeuLeuGlnGlyValArgLeuHisGluLeuLeuPro--- 399
Qy 1196 CATCGCTCCCGCCCAAAAGACAAAGCGCTGACCGCAGCTCGTAACCGAGCTGCAT 1255
Db 1196 CATCGCTCCCGCCCAAAAGACAAAGCGCTGACCGCAGCTCGTAACCGAGCTGCAT 1255
Qy 400 -----MetGlyArgProAspGlySerLeuPheGlySerSerGlyAlaSerLeuHis 416
Db 400 -----MetGlyArgProAspGlySerLeuPheGlySerSerGlyAlaSerLeuHis 416
Qy 1256 GCCAAACCTCTATTGTGGACGCAAAAGCGCATCTTCATCGGCTCATTCACCTCGACCCC 1315
Db 1256 GCCAAACCTCTATTGTGGACGCAAAAGCGCATCTTCATCGGCTCATTCACCTCGACCCC 1315
Qy 417 ThrLysAlaPheValValAspSerAlaGlyPheIleGlySerPheAsnLeuAspPro 436
Db 417 ThrLysAlaPheValValAspSerAlaGlyPheIleGlySerPheAsnLeuAspPro 436
Qy 1316 CGTTCGCGACGGCTCAATACCGAAATCGGCTGCTCATCGAAAGCCCCCAAAATCGCGAAA 1375
Db 1316 CGTTCGCGACGGCTCAATACCGAAATCGGCTGCTCATCGAAAGCCCCCAAAATCGCGAAA 1375
Qy 437 ArgAlaMetAsnLeuAsnThrGluMetGlyLeuPheGluAspArgThrValThrAla 456
Db 437 ArgAlaMetAsnLeuAsnThrGluMetGlyLeuPheGluAspArgThrValThrAla 456
Qy 1376 CAGATGAGCGCCACCCCTC---GCCGATACACACCCCGAATAGCCTACCGGTTACCCCTC 1432
Db 1376 CAGATGAGCGCCACCCCTC---GCCGATACACACCCCGAATAGCCTACCGGTTACCCCTC 1432
Qy 457 GluLeuGluArgLeuTyrAsnHisLysValSerAlaProValSerTyrArgValThrLeu 476
Db 457 GluLeuGluArgLeuTyrAsnHisLysValSerAlaProValSerTyrArgValThrLeu 476
Qy 1433 GACAAACACACCGCTGCAATGCGACGATCCCGCCACCCGCA-----AAACCTACCCG 1486
Db 1433 GACAAACACACCGCTGCAATGCGACGATCCCGCCACCCGCA-----AAACCTACCCG 1486
Qy 477 GluGln---GlyGlyLeuArgTrpHisAspAlaAlaGlnProProAlaValTrpThr 495
Db 477 GluGln---GlyGlyLeuArgTrpHisAspAlaAlaGlnProProAlaValTrpThr 495
Qy 1487 AACGAACCGGAAGCAAACTTTGGAAACGATCGCGCAAAATCTATCCCTCGTCCGCC 1546
Db 1487 AACGAACCGGAAGCAAACTTTGGAAACGATCGCGCAAAATCTATCCCTCGTCCGCC 1546
Qy 496 ArgGluProAlaAlaSerValTrpArgArgGlyAlaAlaLysValLeuGlyTrpLeuPro 515
Db 496 ArgGluProAlaAlaSerValTrpArgArgGlyAlaAlaLysValLeuGlyTrpLeuPro 515
Qy 1547 ATCGAAGGTTTATTA 1561
Db 1547 ATCGAAGGTTTATTA 1561
```

RESULT 29

```
Q8P7A5_XANCP
ID Q8P7A5_XANCP PRELIMINARY; PRT; 520 AA.
AC Q8P7A5;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Cardiolipin synthase.
GN Name=cls;
OS Xanthomonas campestris (pv. campestris).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC Xanthomonadaceae; Xanthomonas.
OX NCBI_TaxID=340;
RN [1]
RP NUCLEOTIDE SEQUENCE
RC STRAIN=ATCC 33913 / NCPB 528;
RX MEDLINE=2022145; PubMed=12024217; DOI=10.1038/417459a;
RA da Silva A.C.R., Reinach F.C., Farah C.S., Furlan L.R.,
RA Quaggio R.B., Monteiro-Vitorello C.B., Van Sluys M.A.,
RA Almeida N.F. Jr., Alves L.M.C., do Amaral A.M., Bertolini M.C.,
RA Camargo L.E.A., Camarotte G., Cannavan F., Cardozo J., Chambergo F.,
```

```
RA Ciapina L.P., Cicarelli R.M.B., Coutinho L.L., Cursino-Santos J.R.,
RA El-Dorry H., Faria J.B., Ferreira A.J.S., Ferreira R.C.C.,
RA Ferro M.I.T., Formighieri E.F., Franco M.C., Greggio C.C., Gruber A.,
RA Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,
RA Locali E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,
RA Martins E.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,
RA Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,
RA Pereira H.A., Rossi A., Sena J.A.D., Silva C., de Souza R.F.,
RA Spinola L.A.F., Takita M.A., Tamura R.E., Teixeira E.C., Tezza R.I.D.,
RA Trindade dos Santos M., Truffi D., Tsai S.M., White F.F.,
RA Setubal J.C., Kitajima J.P.;
RT "Comparison of the genomes of two Xanthomonas pathogens with differing
RT host specificities";
RL Nature 417:459-463(2002).
DR EMBL; AE012383; AAM41978.1; -; Genomic DNA.
DR GO; GO:0003824; F:catalytic activity; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR001736; PLD.
DR Pfam; PF00614; PLDc; 2.
DR SMART; SM00155; PLDc; 2.
DR PROSITE; PS00035; PLD; 2.
KW Complete proteome.
SQ SEQUENCE 520 AA; 57165 MW; 8849CFED240410A3 CRC64;

Alignment Scores:
Pred. No.: 7,16e-39 Length: 520
Score: 843.00 Matches: 196
Percent Similarity: 55.1% Conservative: 71
Best Local Similarity: 40.4% Mismatches: 174
Query Match: 29.6% Indels: 44
DB: 2 Gaps: 12

US-10-665-990A-13 (1-1561) x Q8P7A5_XANCP (1-520)
Qy 170 CTGCAAAATCGCGACACCCCTCATACACGCGGTATCGACATCTACCTCTCGACGAC 229
Db 170 CTGCAAAATCGCGACACCCCTCATACACGCGGTATCGACATCTACCTCTCGACGAC 229
Qy 59 LeuGlnGlnAlaHisAla-----GlyLysThrGlyMetValIleLeuSerAsp 74
Db 59 LeuGlnGlnAlaHisAla-----GlyLysThrGlyMetValIleLeuSerAsp 74
Qy 230 CCCACGAAAGCCCTTGGCGCGCGCTTATCGAATCTGCGGACACAGCCTCGAT 289
Db 230 CCCACGAAAGCCCTTGGCGCGCGCTTATCGAATCTGCGGACACAGCCTCGAT 289
Qy 75 AsnIleAspAlaPheAlaValArgAlaLeuThrAlaArgAlaAlaGlyArgSerLeuAsp 94
Db 75 AsnIleAspAlaPheAlaValArgAlaLeuThrAlaArgAlaAlaGlyArgSerLeuAsp 94
Qy 290 TTGCAATACTACATTTTCGCGCAACGACATTTCCGCGAGGCTGCTGTTCAACCTCATGAT 349
Db 290 TTGCAATACTACATTTTCGCGCAACGACATTTCCGCGAGGCTGCTGTTCAACCTCATGAT 349
Qy 95 LeuGlnTyrTrpIleThrHisAlaAspPheThrGlyAsnLeuLeuHisAsnGluLeuLeu 114
Db 95 LeuGlnTyrTrpIleThrHisAlaAspPheThrGlyAsnLeuLeuHisAsnGluLeuLeu 114
Qy 350 CTTGCGCGCAGAACGCGCGTACGCGTACGCTGCTGTGGACGACAAACACACGCGCGGG 409
Db 350 CTTGCGCGCAGAACGCGCGTACGCGTACGCTGCTGTGGACGACAAACACACGCGCGGG 409
Qy 115 ArgAlaAlaAspArgGlyValArgValArgLeuLeuLeuAspMetAsnIleHisGly 134
Db 115 ArgAlaAlaAspArgGlyValArgValArgLeuLeuLeuAspMetAsnIleHisGly 134
Qy 410 TTGGACGATCTCTGCTGCGCTCGCGCTCGACAGCATCCCAATATCGAAGTGGCGCTTCAAC 469
Db 410 TTGGACGATCTCTGCTGCGCTCGCGCTCGACAGCATCCCAATATCGAAGTGGCGCTTCAAC 469
Qy 135 SerAspSerValLeuAlaAlaLeuAspSerHisProIleLeuGluIleArgLeuPheAsn 154
Db 135 SerAspSerValLeuAlaAlaLeuAspSerHisProIleLeuGluIleArgLeuPheAsn 154
Qy 470 CCCTTCCTCTACGCAAA-----TGGCGCGCACTCGGCTACCTGACGCGCTTCCCC 520
Db 470 CCCTTCCTCTACGCAAA-----TGGCGCGCACTCGGCTACCTGACGCGCTTCCCC 520
Qy 155 ProThrArgAlaAlaArgGluGlyThrLeuMetArgGlyValGluMetValLeuArgMetPhe 174
Db 155 ProThrArgAlaAlaArgGluGlyThrLeuMetArgGlyValGluMetValLeuArgMetPhe 174
Qy 521 CGCCTCAACCGCGCATGACACAAATCTTTACCGCGCAACACCGCGCCACCATCTC 580
Db 521 CGCCTCAACCGCGCATGACACAAATCTTTACCGCGCAACACCGCGCCACCATCTC 580
Qy 175 SerIleAsnArgArgMetHisAsnLysAlaTrpIleAlaAspGlyArgIleAlaValVal 194
Db 175 SerIleAsnArgArgMetHisAsnLysAlaTrpIleAlaAspGlyArgIleAlaValVal 194
Qy 581 GCGCGGACGCAATATCGCGGAGCAATATCTTCAAGTCGGTGGAGGACACCGTTTTCGCGGAC 640
Db 581 GCGCGGACGCAATATCGCGGAGCAATATCTTCAAGTCGGTGGAGGACACCGTTTTCGCGGAC 640
Qy 195 GlyGlyArgAsnValGlyAspGluTyrPheAspAlaAlaArgAspThrAsnPheMetAsp 214
Db 195 GlyGlyArgAsnValGlyAspGluTyrPheAspAlaAlaArgAspThrAsnPheMetAsp 214
Qy 641 CTGGACATCTCTCGCCACCGCGCGTCTGCGCGCAAGTATCGCACGATTCGCGCGTAC 700
Db 641 CTGGACATCTCTCGCCACCGCGCGTCTGCGCGCAAGTATCGCACGATTCGCGCGTAC 700
Qy 215 MetAspAlaAlaLeuIleGlyProThrValArgGlnSerGluGlnValPheAspAlaTyr 234
Db 215 MetAspAlaAlaLeuIleGlyProThrValArgGlnSerGluGlnValPheAspAlaTyr 234
Qy 701 TGGCAAGACCATTCGCGCCCAACCAACCGCACGCGCATCATCCGACG----- 745
Db 701 TGGCAAGACCATTCGCGCCCAACCAACCGCACGCGCATCATCCGACG----- 745
Qy 235 TrpAsnSerProAsnAlaLeuProLeuAlaAlaLeuValThrAlaLysProGlnAlaLeu 254
Db 235 TrpAsnSerProAsnAlaLeuProLeuAlaAlaLeuValThrAlaLysProGlnAlaLeu 254
```


Db 309 GluLysValLeu---ArgLysLysAlaGluAsnTrpLeuMetLysAlaLeuSerGlnVal 327
Qy 974 CTCAAACAGCCGAAAGAGCGTCTATCTGCTTTCACCCCTATTCGTCCTCAAAATCC 1033
Db 328 lIeGluAspAlaGlnLysThrValGlnLysThrProTyrPheValProGlyLysVal 347
Qy 1034 GGCACAGACGCACTGGCAAACTGGTCAGGACGCGATAGACGTTTACCGCTCGACCAAC 1093
Db 348 GlyThrGlnAsnPheSerAsnLeuValSerLysGlyValAspValLysIleLeuThrAsn 367
Qy 1094 TCGCTACAGGACGCGAGTTCGCCGCTCATTCGGCTACGTCAGTCAATACCGAAACCG 1153
Db 368 SerLeuAlaAlaThrAspValAlaLeuValHisGlyGlyTyrValProTyrArgLysMet 387
Qy 1154 CTGCTCAAGCGCGCATCAAACTCTACGAGCTGCAACCCCAACCATCGCTCGCCGCCACA 1213
Db 388 LeuLeuLysSerGlyValLysLeuTyrGluLeuLysProGlyGlyAsnMetHisGlyLeu 407
Qy 1214 AAAGACAAAGCCCTGACCGGAGCTCCGTAACACGCTGCGATGCAAAACCTTCATTGTG 1273
Db 408 Arg-----LeuPheArgSerSerLysAlaSerLeuHisThrLysAlaPheLeuVal 424
Qy 1274 GACGGCAACCGCATCTTCATCGGCTCATTCACCTCGACCCCGCTTCGGACGGCTCAAT 1333
Db 425 AspArgLysThrAlaPheIleGlySerLeuAsnPheAspProArgSerAlaAlaLeuAsn 444
Qy 1334 ACCGAAATGGCGTGGTTCATCGAAGCCCAAAATCGCAGAACAGATGGAGCGCACCTC 1393
Db 445 ThrGluMetGlyIleLeuPheGlyAlaProIleThrThrArgLeuAspLeuPhe 464
Qy 1394 GCC---GATACACACCCGATACGCTTACCGCTTACCGCTTCGACAAACACCGCGCTG 1450
Db 465 AlaGluGluThrThrGlyGluMetSerTyrHisLeuArgLeuAsnAspAsnArgIle 484
Qy 1451 CAATGGCAGATCCCGCACCCGCAAAACCTACCGC-----AACGAAACCGAAGCCAA 1504
Db 485 TyrTrpAspPheIleGluAsnGluLysGlnTyrSerValAsnTyrGluProGluSerAsn 504
Qy 1505 CTTTGGAAACGATCGCGCGAAATCTATCCCTGCTGCGCCATCGAAGGTTTATTA 1561
Db 505 PheTrpArgAlaPheAlaLysIleIleSerTrpLeuProIleGluSerGlnLeu 523

RESULT 32

Q98NZ3_RHILO
ID Q98NZ3_RHILO PRELIMINARY; PRT; 466 AA.
AC Q98NZ3
DT 01-OCT-2001 (TrEMBLrel. 18, Created)
DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DE Phospholipase D-family protein.
GN OrderedLocusNames=mlr9675;
OS Rhizobium loti (Mesorhizobium loti).
OG Plasmid pMdB.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Phyllobacteriaceae; Mesorhizobium.
OX NCBI_TaxID=381;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=MAF303099;
RX MEDLINE=21082930; PubMed=11214968;
RA Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
RA Watanabe A., Idegawa K., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,
RA Mochizuki Y., Nakayama S., Nakazaki N., Shimpo S., Sugimoto M.,
RA Takeuchi C., Yamada M., Tabata S.;
RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium
Mesorhizobium loti.";
RL DNA Res. 7:331-338(2000).
DR EMBL; AP003017; BAB54862.1; -: Genomic DNA.
DR GO; GO:0003824; F:catalytic activity; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR001736; PLD.

DR Pfam; PF00614; PLDC; 2.
DR SMART; SM00155; PLDC; 2.
DR PROSITE; PS00035; PLD; 2.
KW Complete proteome; Plasmid.
SQ SEQUENCE 466 AA; 51793 MW; 4693C68F20F94F3F CRC64;

Alignment Scores:

Pred. No.: 3,57e-37 Length: 466
Score: 812.50 Matches: 197
Percent Similarity: 55.4% Conservative: 63
Best Local Similarity: 42.0% Mismatches: 176
Query Match: 28.5% Indels: 33
DB: 2 Gaps: 12

US-10-665-990A-13 (1-1561) x Q98NZ3_RHILO (1-466)

Qy 200 GGGGTATTCGACATCTACCTGCTCGACAGACCCCAAGCCCTTGGCCGCCGCCGCC 259
Db 16 GlyLysSerGlyLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 35
Qy 260 CTTATCGAATCTCGCAACACAGCCTCGATTGCAATCTACATTTGGCGCAACGACATT 319
Db 36 AlaAlaArgGlyAlaValArgThrLeuAspLeuMetTyrTyrLeuTrpHisAspHis 55
Qy 320 TCCGGCAGGCTGTGTTCAACCTCATCTACCTTCCGCGAGAACCGCGCTCGCGTACGC 379
Db 56 ThrGlyArgLeuLeuLeuGlnValValArgAlaAlaGlnArgGlyValArgValArg 75
Qy 380 CTGCTGTGGACGACAAACACGCGCGGTGGACGATCTCTGCTGCGCCTCGACAGC 439
Db 76 MetLeuLeuAspValAsnProArgLysSerAspAlaAlaTyrLeuAlaLeuAsnAsn 95
Qy 440 CATCCCAATATCGAAGTGCCTTCAACCCCTTCTGCTCTCTACGCAAA-----TGG 490
Db 96 HisProAsnIleGluLeuArgLeuPheAsnProSerGlyIleArgAlaArgGlyLeuMet 115
Qy 491 CGGCGACTCGGCTACCTGACCGACTTCCCGCTCAACCGCCCTCATCGCCATGCACAAATCC 550
Db 116 ArgGlyAlaGluValLeuLeuArgLeuPheAlaLeuLeuThrArgMetHisAsnLysAla 135
Qy 551 TTTACCGCCGACAAACCGCGCCACATATCTCGGCGGACCAATATCGCGACGAATATCTTC 610
Db 136 TrpIleAlaAspSerIleAlaIleValGlyArgAsnValGlyAspAlaTyrPhe 155
Qy 611 AAAGTCGTGAGGACACCGTTTTCGCGACCTGGACATCTCGCCACCGCAGCGTCTGTC 670
Db 156 AspAlaAlaGlu---ThrAsnPheArgAspLeuAspValLeuValLeuGlyProAlaVal 174
Qy 671 GCGCAAGTATCGCAGCACTTCGACCGCTACTGGCGAAGCCATTCCGCCCAACGCCACG 730
Db 175 GlnGlnThrAlaArgIlePheGlnThrPheTrpAla-----CysGlnAspAlaLys 191
Qy 731 CGCATATCTCGCAGCGCAACATCGCAAGGCTTTCAGCACTCGGATAC-----AAC 784
Db 192 ProIleGlyGluLeuGlyAlaAlaProGlySerHisAlaProAsnPheGluArgSer 211
Qy 785 GACGAAACATCCAGACACCGCTCTCGCTACCGGCAACCGTCGCAACAGTCCGCCCTC 844
Db 212 AspAlaGluThrGluSerThrLeuLeu-----SerGlyIle 223
Qy 845 TACCAAAAATAACAGACGGGACGC-----ATCGACTGGCAGAGC 883
Db 224 GlyAspLysGlySerThrAlaGluPheIleAlaAlaSerSerAspValHisTrpValAsp 243
Qy 884 GTCCAAACCGCTGATCAGCGACACCCCTGCAAAAGGACTCGACCGCGACCGCGCAAA 943
Db 244 ---ArgValArgValIleSerAspProGluLys---ValArgGlyTyrPArgProArg 261
Qy 944 CCGCCGATTCGCGGAGGCTGCAAGCGCTCAACAGCCCGCAAAAGGCTCTATCTG 1003
Db 262 GlyTrpLeuMetLysGluLeuLeuProIleIleGlnSerAlaArgLysArgValGluIle 281
Qy 1004 GTTTCACCCCTATTTCGTCCTTACAAAATCCGGCACAGACGACGACTGGCAAAATCGTGCAG 1063


```
QY 692 GACCGCTACTGGGCAAGCCATTCCCGCCACAAAC-----GCCACG 730
DB 228 AspAspTyrTrp-----AsnSerAlaValSerValProValArgSerLeuLeuAlaArg 245
QY 731 CGCATCATCCGAGCGGCAACATCCGGCAAGGCTCTTCAAGCACTCGGATACAAACGAA 790
DB 246 ArgProAsnLysLeuAlaLysLeuArgGluLeuAspAlaLeuProGlnSerGlu--- 264
QY 791 ACATCCAGACACGCGCTCTCGGCTACCGCAAAACCGTCGACAGCTCGCCCTCTACCAA 850
DB 265 -----AlaAlaArgProTyrLeuGluArgValGluSerGln-----TyrGly 278
QY 851 AAA-----ATACAGCGGAGCATCGACTGGCAGAGCGCTCCAAACCGC-----CTG 898
DB 279 ArgAspHisPheLeuMetSerAspArgLeuHisTrp-----ValAspThrAlaAspVal 296
QY 899 ATCAGCGACACCCCTGCAAAAGGATCGACCGGACCGCGCAAAACCGCG-----ATT 952
DB 297 LeuAlaAspProProGluLysAlaAlaGlyLys---ArgArgLysGlyHisAsnPheLeu 315
QY 953 GCCGGAGGCTGCAAGCGCGCTCAACACAGCCGCAAAAGCGTCTATCTGTTTCACCC 1012
DB 316 MetGluSerLeuLeuProLeuMetGlnAlaAlaGlyGluSerLeuHisIleThrSerPro 335
QY 1013 TATTTCTGCTCCCTACAAAATCCGGCACAGACGACCTGGCAAAACCTGGTCGAGCGGCATA 1072
DB 336 TyrPheIleProGlyLysGlnGlyValGluIlePheLeuAspLeuAlaGluArgGlyVal 355
QY 1073 GACGTTACGCTCGTACCAACTCGCTACAGCGACCGCGTTCGCGCGCTCATTCGCGC 1132
DB 356 SerLeuAlaIleLeuThrAsnSerLeuAlaIleThrAspValAlaAlaValHisAlaGly 375
QY 1133 TACGTCAAATACCGAAACCGCTGCTCAAGCGCATCAAACTACGACGTCGACCC 1192
DB 376 TyrAlaArgTyrArgLysProLeuLeuSerGlyGlyValArgLeuHisGluLeuArg--- 394
QY 1193 AACCATGCGCTCCCGCCACAAAAGACAAAGGC-----CTGACCGCGAGCTCCGTA 1243
DB 395 -----SerGlnAlaAspGlnGlySerPheThrLeuArgGlySerGlyGln 409
QY 1244 ACCAGCTGCATGCAAAACCTTCATTGTGACGCGCAAAACGCATCTTCATCGGTCATTC 1303
DB 410 AlaSerLeuHisThrLysAlaPheThrArgAspGlyGluThrGlyTyrIleGlySerLeu 429
QY 1304 AACCTCGACCCCGTTCGCGACGCTCAATACCGAAATGGCGCTGCTCATCGAAAGCCCC 1363
DB 430 AsnPheAspProArgSerAlaSerLeuAsnThrGluMetGlyValValPheAsnSerAla 449
QY 1364 AAAATCGCAGAACAGATGGAGCGCACCTCGCCGATACCAACACCGGAA---TACGCTTAC 1420
DB 450 ProLeuValAlaArgMetAspGluIlePheAlaGluIleArgArgThrMetSerPhe 469
QY 1421 CGCGTTACCTCGCAAAACACACCGCTCGCAATGG-----CAGCATCCGCGCACCGGA 1474
DB 470 GluLeuAspIleAspSerAlaAsnArgIleValTrpMetThrGluGluArgGlyGlnPro 489
QY 1475 AAAACCTACCCGAGACCGACCGCAAACTTTGGAAACGCATCGCGCGCAAAATCCTA 1534
DB 490 LysIleTyrArgArgGluProAspAlaAlaIleSerArgArgIleAlaAlaGlyIleMet 509
QY 1535 TCCCTGCTGCCCATCGAAGGTTTATTA 1561
DB 510 ArgValLeuProLeuGluSerGlnLeu 518
RESULT 34
Q984B3_RHILO
ID Q984B3_RHILO PRELIMINARY; PRT; 466 AA.
AC Q984B3;
DT 01-OCT-2001 (TrEMBLrel. 18, Created)
DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Cardiolipin synthetase family protein.
```

```
OrderedLocusNames=mlr8077;
OS Rhizobium loti (Mesorhizobium loti).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OX Phyllobacteriaceae; Mesorhizobium.
RN NCBI_TaxID=381;
RP NCU
RC STRAIN=MAFF303099;
RX MEDLINE=21082930; PubMed=11214968;
RA Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
RA Watanabe A., Iidesawa K., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,
RA Mochizuki Y., Nakayama S., Nakazaki N., Shimpo S., Sugimoto M.,
RA Takeuchi C., Yamada M., Tabata S.;
RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium
RT Mesorhizobium loti";
RL DNA Res. 7:331-338(2000).
DR EMBL; BA000112; BAB53717.1; -; Genomic DNA.
DR GO; GO:0003824; F:catalytic activity; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR001736; PLD.
DR Pfam; PF00614; PLD; 2.
DR SMART; SM00155; PLD; 2.
DR PROSITE; PS00035; PLD; 2.
KW Complete proteome.
SQ SEQUENCE 466 AA; 52117 MW; 693301E508A0A788 CRC64;
Alignment Scores:
Pred. No.: 9.99e-37 Length: 466
Score: 804.50 Matches: 193
Percent Similarity: 53.9% Conservative: 61
Best Local Similarity: 41.0% Mismatches: 180
Query Match: 28.2% Indels: 37
DB: 2 Gaps: 10
US-10-665-990A-13 (1-1561) x Q984B3_RHILO (1-466)
```

```
QY 200 GGGCTATCCGACATCTACCTGCTCGACGACCCCGACGAGCCCTTGGCGCGCGCCGCC 259
DB 16 GlyLysSerGlyLeuLeuLeuIleSerAspAsnTyrAspAlaPheAlaAlaArgValLeu 35
QY 260 CTTATCGAATCTGCGCAACACAGCCTCGATTTGCAATACTACATTTGGCGCAACGACATT 319
DB 36 AlaAlaArgGlyAlaAlaArgThrLeuAspLeuMetTyrTyrLeuTrpHisAspAspHis 55
QY 320 TCGCGAGGCTGCTGTTCAACCTCATGTCTCCCGCAGAACCGCGGTCGCGGTACGC 379
DB 56 ThrGlyArgLeuLeuLeuGlnValValArgAlaAlaGlnArgGlyValArgValArg 75
QY 380 CTGCTGTGTGACGACAAACACGCGCGGTGTGGACGATCTCTCTGCTCGCCCTCGACAGC 439
DB 76 MetLeuLeuAspAspValAsnProArgLysSerAspAlaIleTyrLeuAlaLeuSerAsn 95
QY 440 CATCCCAATATCGAAGTGGCGCTGTTCAACCCCTTCCTCTACGCAAA-----TGG 490
DB 96 HisProAsnIleGluLeuLysLeuPheAsnProSerGlyIleArgAlaArgGlyLeuMet 115
QY 491 CGCGCATCTCGGTACTGACGAGCTTCCCGCGCTCAACCGCGCATGCGCAACAATCC 550
DB 116 ArgGlyAlaGluValLeuLeuArgLeuPheAlaLeuThrArgArgMetHisAsnLysAla 135
QY 551 TTTTACCGCCGACAAACCGCGCACCATCTCGCGGAGCAATATCGCGGACGAAATCTTC 610
DB 136 TrpIleAlaAspAsnIleAlaIleValGlyArgAsnValGlyAspAlaTyrPhe 155
QY 611 AAAAGTCGGTGAAGACACCGGTTTTTCGCGACTCGACATCTCGACACCGCGGAGCGTGC 670
DB 156 AspAlaAlaGlu---ThrAsnPheArgAspLeuAspMetLeuLeuLeuGlyProAlaVal 174
QY 671 GCGGAAGTATCGCAGACCTTCGACCGCTACTGGCAAGCCATTCCGCGCCACAGCCACG 730
DB 175 GlnGlnThrAlaGlnIlePheGlnThrPheTrpValCysGlnAspAla----- 190
```

```
QY 731 CGCATCATCCGCGGCAACATCGCGAAGGTCTTCAAGCACTCGGATACACACGAA 790
Db 191 -----LysProfileAlaGluLeuGlyAlaAlaPro 201
QY 791 ACATCCAGACACGCGCTCTCGGTACCGCGAAACCGTCAAGACATCGCCCTTACCAA 850
Db 202 GlySerHisAlaProTyrPheGluGlyArgGluGluLysThrGluSerThrLeuLeuSer 221
QY 851 AAAATACAG-----ACGGGAGCATCGACTGG 877
Db 222 GlyIleArgAspLysGlySerIleAlaGluPheIleSerAlaSerSerAsnValHisTrp 241
QY 878 CAGAGCGTCCAAACCGGCTGATAGCGACACCCCTGCAAAAGAGCTCGACCGCACGC 937
Db 242 ValGlu---ArgValArgValIleSerAspProGluLys---ValArgGlyTrpArg 259
QY 938 CGCAACACCGCATCGCGGAGGTGCAAGACGCGCTCAACACGCCGCAAAAAGCGTC 997
Db 260 ProArgSerTrpLeuMetLysGluLeuLeuProIleIleGlnSerAlaArgLysArgVal 279
QY 998 TATCTGTTTCCCTATTTCCTACAAATCCGGACAGACGCACTGGCAAAACTG 1057
Db 280 GluIleValSerProTyrPheIleProGlyLysLysGlySerLysIleLeuGlyAspLeu 299
QY 1058 GTGCGAGCGCATAGACGTTACCTCTGACCAACTCGCTACAGCGACGACGCTGCC 1117
Db 300 ValAspAspGlyValGlnValAlaValLeuThrAsnSerLeuAlaAlaThrAspValAla 319
QY 1118 GCGGTCCATTCGGGTCAATACCGAAACCGCTGCTCAAAAGCGGATCAAACTC 1177
Db 320 AlaValHisGlyAlaTyrAlaAsnTyrArgLysArgLeuLeuArgMetGlyValGlnLeu 339
QY 1178 TAGAGCTGCAACCCAAACATGCCGTCCCGGCCCAAAAAGACAAAGGCTGACCGGACG 1237
Db 340 PheGluLeuGlnProPheSerArgGlnProLysIle-----SerValPheGlySer 356
QY 1238 TCCGTAAACGAGCTCATGCCCAAAACCTTCATTGTGGAGCGGCAACGATCTTCATCGC 1297
Db 357 LysGlyAlaSerLeuHisThrLysAlaPheSerValAspAsnArgIleGlyPheValGly 376
QY 1298 TCATTCAACCTCGACCCCGTTCGCGACGGCTCAATACCGAATGGGGTTCGTATCGAA 1357
Db 377 SerPheAsnPheAspProArgSerValSerLeuAsnSerGluMetGlyValLeuPheGlu 396
QY 1358 AGCCCCAAA---ATCGCAAGACAGATGGAGCGCACCTCGCGATACACACCGCAATAC 1414
Db 397 AspGluAsnLeuValAlaGluLeuArgHisArgPheLysSerGluIleAlaProGluAla 416
QY 1415 GCCTACCGGTACCTCGACAAACACACCGCTGCAATGGCAC-----GATCCCGCC 1468
Db 417 SerTyrArgLeuGluLeu---LysAsnGluValLeuArgTrpPheIleGlySerAspGluGly 435
QY 1469 ACCCGAAACACCTACCCGAACCAACCGAAGCGCAAACTTTGAAACGGATCCGCCGCAAAA 1528
Db 436 ArgLeuGlnThrTyrThrArgGluProGluAlaAlaTrpPheArgArgIleLeuAlaAla 455
QY 1529 ATCTATCCCTGCTGCCCATCGAAGGTTTATTA 1561
Db 456 LeuValArgHisLeuProIleGluSerGlnLeu 466
RESULT 35
Q6FB53 ACIAD PRELIMINARY; PRT; 520 AA.
AC Q6FB53;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Putative phospholipase D protein.
GN OrderedLocustNames=ACIAD1881;
OS Acinetobacter sp. (strain ADP1).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Moraxellaceae; Acinetobacter.
OX NCBI_TaxID=62977;
```

```
[1] NUCLEOTIDE SEQUENCE.
RP STRAIN=ADP1;
RC PubMed=15514110; DOI=10.1093/nar/gkh910;
RA Barbe V., Vallenet D., Fonknechten N., Kreimeyer A., Oztas S.,
Labarre L., Cruveiller S., Robert C., Duprat S., Wincker P.,
Ornston L.N., Weissenbach J., Marliere P., Cohen G.N., Medigue C.;
RT "Unique features revealed by the genome sequence of Acinetobacter sp.
ADP1, a versatile and naturally transformation competent bacterium.";
RL Nucleic Acids Res. 32:5766-5779(2004).
DR EMBL; CR543861; CAG68710.1; -, Genomic DNA.
DR GO; GO:0016740; P:transferase activity; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR001736; PLD.
DR Pfam; PF00614; PLDC; 2.
DR SMART; SM00155; PLDC; 2.
DR PROSITE; PS00035; PLD; 2.
KW Complete proteome.
SQ SEQUENCE 520 AA; 60385 MW; 1DD312966EA749A0 CRC64;

Alignment Scores:
Pred. No.: 3.9e-36 Length: 520
Score: 794.00 Matches: 186
Percent Similarity: 53.1% Conservative: 76
Best Local Similarity: 37.7% Mismatches: 177
Query Match: 27.8% Indels: 54
Db: 2 Gaps: 11

US-10-665-990A-13 (1-1561) x Q6FB53_ACIAD (1-520)
QY 182 CACACCCCTCATAAC-----AACGGGTATCCGACATCTACCTGCTCGACGACCCCCAC 235
Db 49 HisThrAspGluAsnLeuGlnLysGlyLeuThrAlaTyrIleProLeuAspAspAlaPhe 68
QY 236 GAAGCCCTTGGCGCGCGCGCCCTTATCGAAATCTGCGAACACAGCTCGATTGCAA 295
Db 69 IleSerIleAlaSerArgIleTyrLeuIleArgHisAlaLysHisGlnIleAspLeuGln 88
QY 296 TACTATATTTGGCGCAACGACATTTCCGGGAGGTGCTGTTCACCTCATCTACCTTCCG 355
Db 89 TyrTyrIleTrpHisAsnAspPheValGlyGlnLeuIleLeuAsnGluLeuLysAla 108
QY 356 GCAGAACGCGGTGCGGTACGCTGTGTGGACGACAAACACACGCGGGTGGAC 415
Db 109 AlaAspArgGlyValLysIleArgLeuLeuIleAspGlnAsnGlyThrArgLeuAsp 128
QY 416 GATCTCTGCTCGCCCTCGACAGCATCCCAATATCGAAGTCGCGCTTCAACCCCTTC 475
Db 129 SerThrPheGlnAlaLeuGlnHisProLeuPheGluIleLysValPheAsnProTyr 148
QY 476 GTCCTACGCAATGGCGGCACTCGGTACCTGACCGACTTCCCGCGCTCAACCGCGC 535
Db 149 LysPheArgLysLeuArgPheValAspPheLeuPheArgProAsnGlnIleAsnHisArg 168
QY 536 ATGCACACAACATCTTTTACCGCGACAAACCGCGCACCATATCGCGGCGACGCAATATC 595
Db 169 MethHisAsnLysLeuIleAlaAspAsnThrIleAlaValThrGlyValArgAsnIle 188
QY 596 GCGCGCAATACTTCAAAAGTCGGTAGGACACCGGTTTTCGCGACCTCGACATCTCGCC 655
Db 189 SerSerGluTyrPheAspAlaSerTyrGlnPheGlnPheThrAspLeuAspIleLeuPhe 208
QY 656 ACCGCGACGCTGTCGGCGAAGTATCGCACACTTCGACACTCTACTGGCGAGCCATTCC 715
Db 209 TyrGlyThrAlaAlaSerGlnAlaGluLysValPheGlnThrPheTrpAspSerProLeu 228
QY 716 GCCCAACACCGCACCGCATCATCGCAGCGCGCAACATCGGCAAGGCTCTCAACGACTC 775
Db 229 SerValThrThrGlnLysLeuIleThrLysGlyGlnProAsnAspLeuLeuGlnLeuArg 248
QY 776 GGATACAAACGACGAAACATCCAGACACGCGCTCTCGCGCTACCGCGAAACCGTCGAACAG 835
Db 249 Ser-----LysTyrLysGlu---IleGluArg 256
```

```
QY 836 -----TCGCCCCTCTACCAAAATAACAGACGGGACGCATCGAC----- 874
DB 257 ValAspThrProThrgluAspLysIleAsnGluAlaGlnIleGluLeuAsnGluAlaLeu 276
QY 875 -----TGGCAGAGCTCCAAACCCCGCTGATCAGGACACCCCT 913
DB 277 LysGlnArgTyrSerValLysTrp-----AlaLysAlaHisPheValAlaAspSerPro 294
QY 914 GCAAAAGACCTCGACCGCGACCGCAACCGCGATTGCGGGAGGCTCAAGACGCG 973
DB 295 LysLysIleHisGlyHisAlaIleGlyAsnGluLeuIleTyrAsnGlnMetPheSerIle 314
QY 974 CTCAAACAGCCGCAAAAGCGTCTATCGTTTTCACCTATTTCGTCCTTACAAATCC 1033
DB 315 MetGlyHisProGlnLysHisLeuGluLeuValSerAlaTyrPheValProThrGlnAla 334
QY 1034 GGCACAGGCACTGGCAAACTGTGTGAGACGGCAGTAGACGTTACGCTCAGCAAC 1093
DB 335 GlyThrGlnTyrLeuSerAsnLeuSerGlnGlnAspValLysIleArgValLeuThrAsn 354
QY 1094 TCGCTACAGGGACCGAGTTCGCGCTGCTCATTCCGCTACGTCAGTCAATACCGAAACCG 1153
DB 355 SerPheAlaAlaAsnAspValAlaIleValHisAlaPheTyrSerGlnTyrArgLysGlu 374
QY 1154 CTGCTAAAGCCGCATCAACTCTACGAGCTGCAACCC----- 1192
DB 375 LeuIleLysAsnGlyIleGlnLeuTyrGluPheLysProIleIleAspArgLysGluPro 394
QY 1193 -----AACATGCGCTCCCGCCCAAAAGACAAAGCGCTGACC 1231
DB 395 ThrTrpTyrGluLysMetThrGlyArgValIleProAla-----Lys 408
QY 1232 GGCAGCTCCGTAACAGCCTGCATGCCAAACCTTCATTGTGGAGCGCAACGCATCTTC 1291
DB 409 GlyLysLysSerSerLeuHisAlaLysPheAspValAspGlyLys---ValPhe 427
QY 1292 ATCGGCTCATCAACCTCGACCCCGTTCGCGACGGCTCAATACCGAAATGGCGTCTC 1351
DB 428 IleGlySerPheAsnPheAspProArgSerAlaTyrLeuAsnThrGluValGlyLeuVal 447
QY 1352 ATCGAAAGCCCAAAATCGCAGAACACAGATGAGCGCACCTCGCCGATACACACCGGAA 1411
DB 448 ValGluSerGluAspLeuGlnAsnGluIleThrLysValLeuAspGluTyrLeuProThr 467
QY 1412 TAGCCCTACCGCTTACCTCGACAAACACACCGCTGCAATGGCAGCAT---CCGCGC 1468
DB 468 IleAlaTyrGlnLeuLysLeuAspLysAsnGlyAsnLeuValTrpLeuAspHisAsnAla 487
QY 1469 ACCCGAAACACC-----TACCCGAAACGAAACCCGAAACCTTTTGGAAACGCATCGCC 1522
DB 488 AspGlyGlnThrIleGluTyrHisHisAspProGluThrThrArgPheGlnArgPheMet 507
QY 1523 GCAAAATCTATCCCTGCTGCCCATCGAAGTTTATTA 1561
DB 508 MetLysAlaValSerTyrLeuProIleGluTrpMetMet 520

RESULT 36
O9HTP4_PSEAE
ID O9HTP4_PSEAE PRELIMINARY; PRT; 529 AA.
AC Q9HTP4;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein.
GN OrderedLocNames=PA5310;
OS Pseudomonas aeruginosa.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OX Pseudomonadaceae; Pseudomonas.
NCBI_TaxID=287;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=ATCC 15692 / PA01;
```

```
RX MEDLINE=20437337; PubMed=10984043; DOI=10.1038/35023079;
RA Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warren P.,
RA Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
RA Garber R.L., Goltry L., Tolentino E., Westbrook-Wadman S., Yuan Y.,
RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
RA Reizer J., Saier M.H. Jr., Hancock R.E.W., Lory S., Olson M.V.;
RT "Complete genome sequence of Pseudomonas aeruginosa PA01, an
RT opportunistic pathogen.";
RL Nature 406:959-964(2000).
DR EMBL; AE004943; AAG08695.1; -; Genomic_DNA.
DR PIR; F82983; F82983.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR GO; GO:0008654; P:phospholipid biosynthesis; IEA.
DR InterPro; IPR001736; FLD.
DR Pfam; PF00614; PLDC; 2.
DR SMART; SM00155; PLDC; 2.
DR PROSITE; PS00035; PLD; 2.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 529 AA; 59199 MW; DABF5ED5881E63DE CRC64;

Alignment Scores:
Pred. No.: 5,05e-36 Length: 529
Score: 792.00 Matches: 200
Percent Similarity: 52.5% Conservative: 86
Best Local Similarity: 36.7% Mismatches: 189
Query Match: 27.8% Indels: 70
DB: 2 Gaps: 16

US-10-665-990A-13 (1-1561) x Q9HTP4_PSEAE (1-529)
QY 47 CGCAGCTCATTTCCCTTTTATCGCTCTCTCTGTTCTATGTTCTTCATGTTGCCCCCA 106
DB 15 ArgArgPheLeuLeuLeuAlaLeuLeuAlaLeuLeuAlaLeuLeuAlaLeuLeuAlaLeu 34
QY 107 -----CTGGAAGAACCGACGGAAGCCGT 130
DB 35 GlnProSerSerAlaLeuProAlaGluGlyThrTrpLeuAlaArgGlnAlaGlu----- 52
QY 131 CATTTCAATACTTCCAAACCTGTCTCTGGACAAACATCTGCAATTCGGCACACCCCT 190
DB 53 -----IleGlnGlyArgAspHisPro 59
QY 191 CATAACACAGGGCTATCCGACATCTACCTGCTCGACGACCCCGACGAGCCCTTGGCGC 250
DB 60 -----GlyGlnSerGlyPheHisLeuLeuSerAlaSerGluAspAlaPheValAla 76
QY 251 CGCGCCGCTTATCAATCTGCGAAACACAGCTCGATTGCAATCTACTACTATTTGGCGC 310
DB 77 ArgAlaAlaLeuIleArgAlaAlaGlnArgSerLeuAspIleGlnTyrTyrIleValHis 96
QY 311 AACGACATTTCCGGCAGGCTGTGTTCAACCTCATGTACTTCCCGCAGAACCGCGGTG 370
DB 97 AspGlyLeuThrThrArgAlaLeuAlaTyrGluLeuLysAlaAlaAspArgGlyVal 116
QY 371 CGCTAGCGCTGCTGTGGACGACAAACACACGCGGGGTGGACCATCTCTGCTGCTGCC 430
DB 117 ArgValArgIleLeuIleAspSerThrAlaSerAspGlyTrpAspTyrGluIleGlyVal 136
QY 431 CTCGACAGGCATCCCAATATCGAAGTGGCGCTGTTCAACCCCTTCGCTCTACGCAAA--- 487
DB 137 LeuSerAlaHisProAsnIleGlnValArgLeuPheAsnProLeuHisLeuGlyArgAla 156
QY 488 -----TGGCGCGCATCTCGGCTACTCGACGACTTCCCGCCCTCAACCCCGCATG 538
DB 157 ThrGlyIleThrArgGlyValGlyArgLeuPheAsnLeuSerGlnGlnHisArgMet 176
QY 539 CACAACAAATCTTTTACCGCCGACCAACCGCGCCACCATCTCGCGCGGAGCATATCGGC 598
DB 177 HisAsnLeuLeuTrpLeuAlaAspGlyThrAlaAlaIleValGlyArgAsnLeuGly 196
QY 599 GACGAATACTTCAAAGTCGTGAGGACACCGTTCGCGCGACCTGGACATCTCTCGCCACC 658
```

```
Db 197 AspGluTyrPheAsnAlaLysProGluMetAsnPheThrAspLeuAspLeuGlyVal 216
|||||
::: ||| |||||:::
Qy 659 GGCAGCGTCGCGGAAGTAGTCGACACGACTTCGACCGCTACTCGGCGAAGCCATTCGCC 718
::: |||||::: |||||::: |||||:::
Db 217 GlyProIleAlaAsnGlnLeuSerHisPheAspGlnTyrTrpAsnSerAlaIleSer 236
|||
::: |||||::: |||||::: |||||:::
Qy 719 CACAACGCCAGC-----CGCATCATCCGACGCGGCAACATCGCAAGGCTTT 766
|||
::: |||||::: |||||:::
Db 237 ArgProIleGluAspPheLeuTrpArgAlaProTyrProGlyGluLeu----- 252
|||
::: |||||::: |||||:::
Qy 767 CAAGCACTCGGATACACGACGAAATCCAGACGCGCTCTCGGTAC-----CGC 820
|||
::: |||||::: |||||:::
Db 253 -----GluSerAlaArgArgLysLeuGlnArgTyrLeuArgLys 265
|||
::: |||||::: |||||:::
Qy 821 GAAACCGTCGACAGTCG-----CCCTCTACCAAAAATACAGACGGGAGCG 868
|||
::: |||||::: |||||:::
Db 266 GluSerValLysGluSerGlyTyrIleArgHisLeuPheAspArgLysGlnProArg 285
|||
::: |||||::: |||||:::
Qy 869 ATC-----GACTGCCAGC-----GTCCAAACCGCGCTGATCAGGACACCCCT 913
|||
::: |||||::: |||||:::
Db 286 LeuGlyAsnTrpLeuGluAsnLeuThrTrpAlaArgAlaGluAlaIleTrpAspAlaPro 305
|||
::: |||||::: |||||:::
Qy 914 GCAAAAGGACTCGACCGCGCGCGCAAAACGCG-----ATTGCGGAGGCTGCAA 967
|||
::: |||||::: |||||:::
Db 306 LeuLysValLeuSerArgGlyGluProAspProHisLeuLeuLeuSerProHisLeuAla 325
|||
::: |||||::: |||||:::
Qy 968 GACGCGCTCAACACGCGCGGAAAAAGCGTATCTGTGTTTCCACCTATTTCGTCCTACA 1027
|||
::: |||||::: |||||:::
Db 326 GlyLeuPheLysGlyValGlnLysGluLeuIleValSerAlaTyrPheValProAla 345
|||
::: |||||::: |||||:::
Qy 1028 AAATCCGGCAGACGACCTGCAAACTGGTGAGGAGCGCATAGAGTTACCGTCTCTG 1087
|||
::: |||||::: |||||:::
Db 346 LysAspGlyLeuAsnTyrLeuThrGlyLysAlaAspSerGlyValArgValArgLeuLeu 365
|||
::: |||||::: |||||:::
Qy 1088 ACCAACTCGCTACGCGCAGCGCTGCGCGCTTCATTCGCTCGCTCAAAATACCGA 1147
|||
::: |||||::: |||||:::
Db 366 ThrAsnSerLeuGluAlaThrAspValProAlaValHisAlaGlyTyrAlaProTyrArg 385
|||
::: |||||::: |||||:::
Qy 1148 AAACCGCTGCTCAAGCGCGCATCAAACTTACGAGCTGCAACCCACCGCTCGCTCCC 1207
|||
::: |||||::: |||||:::
Db 386 MetAlaLeuLeuGluHisGlyValLysLeuTyrGluLeuArgAlaAsnProAspGlnPro 405
|||
::: |||||::: |||||:::
Qy 1208 GCCACA---AAGACAAAGGCTGACCGCGCTCGTAACCGAGCTGCATCCGCAAAACC 1264
|||
::: |||||::: |||||:::
Db 406 LeuSerGlyAlaProTrpArgLeuHisGlySerSerAlaSerLeuHisSerLysAla 425
|||
::: |||||::: |||||:::
Qy 1265 TTCATTGTGGCGCAACGCACTTCATCGGCTCATCAACCTCGACCCCGGTTCGCA 1324
|||
::: |||||::: |||||:::
Db 426 MetValPheAspArgLysValPheIleGlySerPheAsnPheAspProArgSerIle 445
|||
::: |||||::: |||||:::
Qy 1325 CGGCTCAATACGAAATGGCGCTGCTCATCGAAAGCCCAAAATCGAGAACAGATGGAG 1384
|||
::: |||||::: |||||:::
Db 446 LeuTrpAsnThrGluValGlyValIleValAspSerProLeuLeuAlaGluGlnValArg 465
|||
::: |||||::: |||||:::
Qy 1385 CGC---ACCTCGCGCATACACACCCGNAATAGCGCTACCGGTTACCTCGACAAACAC 1441
|||
::: |||||::: |||||:::
Db 466 GlnLeuAlaLeuGluGlyMetAlaProSerValSerTyrGlnValArgIleAspArgSer 485
|||
::: |||||::: |||||:::
Qy 1442 AAC-----CGCTGCAATGGCAGCAT-----CCGCGCACCCGAAAAACCTACCCG 1486
|||
::: |||||::: |||||:::
Db 486 GlySerArgProLysLeuValTrpIleAspGluArgAspGlyValArgAlaGlnValLeuArg 505
|||
::: |||||::: |||||:::
Qy 1487 AACGAACCGCAAGCAAACTTTGGAACGCAATCGCGCAAAAATCCTATCCCTGTCGCC 1546
|||
::: |||||::: |||||:::
Db 506 HisGluPro---GlySerLeuTrpArgArgLeuAsnAlaTrpValAlaGlyMetIleGly 524
|||
::: |||||::: |||||:::
Qy 1547 ATCAGAGGTTTATTA 1561
|||
::: |||||::: |||||:::
Db 525 LeuGluLysMetLeu 529
|||
::: |||||::: |||||:::
RESULT 37
Q62E43_BURMA
```

```
Q62E43_BURMA PRELIMINARY; PRT; 550 AA.
Q62E43:
25-OCT-2004 (TrEMBLrel. 28, Created)
25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
Putative phospholipase D.
OrderedLocusNames=BMAA0208;
Burkholderia mallei (Pseudomonas mallei).
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Burkholderiaceae; Burkholderia.
OX NCBI_TaxID=13373;
RN [1]
NUCLEOTIDE SEQUENCE.
RC STRAIN=ATCC 23344;
RX PubMed=15377793; DOI=10.1073/pnas.0403306101;
RA Nieman W.C., Deshazer D., Kim H.S., Tettelin H., Nelson K.E.,
RA Feldblyum T.V., Ulrich R.L., Romning C.M., Brinkac L.M.,
RA Daugherty S.C., Davidlsen T.D., DeBoy R.T., Dimitrov G., Dodson R.J.,
RA Durkin A.S., Gwinn M.L., Haft D.H., Khouri H.M., Kolonay J.F.,
RA Madupu R., Mohammed Y., Nelson W.C., Radune D., Romero C.M.,
RA Sarria S., Selengut J., Shambhlin C., Sullivan S.A., White O., Yu Y.,
RA Zafar N., Zhou L., Fraser C.M.;
RT "Structural flexibility in the Burkholderia mallei genome.";
RL Proc. Natl. Acad. Sci. U.S.A. 101:14246-14251(2004).
EMBL: CP000011; AAU46047.1; -, Genomic_DNA.
TIGR: BMAA0208; -.
DR GO: GO:0003824; F:catalytic activity; IEA.
DR GO: GO:0008152; P:metabolism; IEA.
DR InterPro: IPR001736; PLD.
DR Pfam: PF00614; PLDC; 2.
DR SMART: SM00155; PLDC; 2.
DR PROSITE: PS00035; PLD; 2.
DR Complete proteome.
SQ SEQUENCE 550 AA; 59587 MW; 83820C127A80DBA4 CRC64;

Alignment Scores:
Pred. No.: 5,13e-35 Length: 550
Score: 774.00 Matches: 195
Percent Similarity: 51.1% Conservative: 81
Best Local Similarity: 36.1% Mismatches: 216
Query Match: 27.1% Indels: 48
DB: Gaps: 13

US-10-665-990A-13 (1-1561) x Q62E43_BURMA (1-550)
Qy 47 CGCAGCTCATTTCCCTTTTATGCTCTTCTCTGTTTCATGTTCTATGTTGCTCCCA 106
|||||
::: |||||::: |||||::: |||||:::
Db 24 ArgAlaLeuValLeuCysAlaAlaLeuSerGlyCysAlaThrHisProProAla 43
|||
::: |||||::: |||||::: |||||:::
Qy 107 CTGGAA---GAACGCGGAAAGCGGTCAATTTCAATCTCCAAACCTGCTCTCTCGAC 163
|||||
::: |||||::: |||||::: |||||:::
Db 44 ThrThrLeuGluAlaGlnValSerHisAlaLeuProProAspAlaSerThrProLeuAla 63
|||
::: |||||::: |||||::: |||||:::
Qy 164 AACATCTCTGCAAAATCCGCGCACACCCCTCATAAACCGGCTATCCGACATCTACCTGCTC 223
|||
::: |||||::: |||||::: |||||:::
Db 64 AspAlaLeuAlaValGlnAlaArgAlaHis---ProGlyGluSerGlyPheValValLeu 82
|||
::: |||||::: |||||::: |||||:::
Qy 224 GACGACCCCAACGAGCCCTTTCGCGCCGCCCTTATCGAATCTGCGCAACACAGC 283
|||||
::: |||||::: |||||::: |||||:::
Db 83 ProArgGlyAspGluAlaLeuGlnMetArgIleAlaValAlaArgAlaThrLysThr 102
|||
::: |||||::: |||||::: |||||:::
Qy 284 CTCGATTTGCAATACATTTGGCGCAACGACATTTCCGGCAGCTGCTGTTCAACCTC 343
|||||
::: |||||::: |||||::: |||||:::
Db 103 LeuAspIleGlnTyrTyrIleAlaGluAspThrThrGlyLysLeuLeuGlyAla 122
|||
::: |||||::: |||||::: |||||:::
Qy 344 ATGTACCTTTCGCGAAGCGCGGTGCGCTAGCTGCTGCTGTTGGACGACCAACAACG 403
|||||
::: |||||::: |||||::: |||||:::
Db 123 AlaLeuTyrAlaAlaAspArgGlyValArgValArgMetLeuValAspAlaLeuAsnPhe 142
|||
::: |||||::: |||||::: |||||:::
Qy 404 CGCGGTTGGACGATCTCTCTGCTCCCTCGACAGCCATCCCAATATCGAAGTGGCGCTG 463
|||||
::: |||||::: |||||::: |||||:::
Db 143 LysAspIleAspLysLeuMetAlaLeuAspAlaHisAlaAsnLeuGluValArgVal 162
|||||
::: |||||::: |||||::: |||||:::
```

```
QY 464 TTCACCCCTTCTGCTCTACGCAAAATGG-----CGCGCACTCGGTACTCGTACC 511
Db PheAsnProPheGlyAlaProArgLeuGlyMetPheAlaArgThrAlaAsnValPheThr 182
QY 512 GACATTCCTCCCGCTCAACCGCGCATGACACAAATCCTTTACCGCGGACAAACGGCC 571
Db ArgIleAspAsnPheThrArgArgMethHisAsnYsAlaMetIleSerAspAsnGlnIle 202
QY 572 ACCATCTACTGGCGGCAATATCGCGCAGCAATACTTCAAAATCGGTGAGGACACCGTT 631
Db AlaIleValGlyAlaAsnLeuGlyAspGluTyrPheAsnAlaSerProThrLeuGln 222
QY 632 TTCGCGGACCTTGACATCTCGCCACCGCGAGCGTGTGCGGAGTATCGACAGATTC 691
Db PheArgAspLeuAspValLeuAlaGlyProValThrArgAlaValSerAlaSerPhe 242
QY 692 GACCGCTACTGGCAAGCATTCGCGCCACACCGCCAGCGCATCTCCGAGCGGCAC 751
Db AspAlaTyrTyrPheSerAlaLeuThrTyrProLeuProAlaLeuAsnArgArgTyr 262
QY 752 ATCGCGCAAGGTCTTCAAGCACTCGGATACACGAGCAAAACATCCAGACCGCTCTG 811
Db AspAlaTyrAspLeuAspAla-----AlaArgAspAlaLeuArg 275
QY 812 CGCTAC-----CGCGAAACCGTCGACAGTCCGCCCTC 844
Db AlaHisTyrAlaAsnAlaThrProTyrAsnAlaLysProLeuAsnAlaThrProLeu 295
QY 845 TACCAGAAATACAGCGGAGCATCGAC-----TGGCAGAGCTCCAAACCGCGCTG 898
Db AlaAlaGlnIleAlaArgAsnGluLeuGlyLeuValTyrAlaSerAlaGlu-----Phe 313
QY 899 ATCAGCGACACCCCTGCAAAA---GGACTCGACCGCGACCGCCGCAACCGCGATTGCC 955
Db ThrAlaAspSerProGluLysIleAlaAlaProAspAspSerTyrLysSerProMet 333
QY 956 GGGAGGTGCAAGCGCTCAACAGCCGCAAAAGCGCTATCTGTTTTCACCCCTAT 1015
Db GlnArgLeuPheAlaLeuThrArgAspAlaGlnArgGluPheLeuValLeuSerProTyr 353
QY 1016 TTCCTCTTCAAAATCCGCGACAGCGCTGCAAAATCGGTGCGAGGAGGATAGAC 1075
Db PheValProHisAspAlaGlyValAsnAlaLeuGlyArgLeuThrAlaArgValArg 373
QY 1076 GTTACCTCTCTGCAACCTCGCTACAGCGCGACCGTTCGCGCTCATTCGCGGTAC 1135
Db ValAlaIleLeuThrAsnSerLeuAlaAlaThrAspAlaIleAlaValGlnAlaGlyTyr 393
QY 1136 GTCAAATACGAAACCGCTCTCAAGCGGCAATCAAACTCTACGAGTGCACCCCAAC 1195
Db AlaProTyrArgValProMetLeuGluArgGlyValGluLeuTyrGluTyrLysProAsp 413
QY 1196 CATCCCTCCCGCCCAAAAAGCAAAAGGCTGACCGGACG---TCGTAACAGGCTG 1252
Db -----ProGlyArgSerArgIleGlyMetLeuGlySerArgSerAlaSerLeu 430
QY 1253 CATGCCAAAACCTTCAATGTGGAGCGCAACGCGATCTTCAATCGCTCATTCACCTCGAC 1312
Db HisAlaLysAlaTyrValIleAspArgLysIleLeuValIleGlySerMetAsnLeuAsp 450
QY 1313 CCCCGTTCCGACGGCTCAATACCGAAATGGGGTCTGTCATCGAAAGCCCAAAATCGCA 1372
Db ProArgSerAlaHisLeuAsnThrGluLeuAlaLeuValIleHisSerProArgLeuAla 470
QY 1373 GAACAGATGGAGCGCCTCGCGATACACACA---CCGAAATACGCTTACCGGTTACC 1429
Db AsnGluValAlaAlaAsnLeuPheAspGluValThrLysProThrIleSerTyrArgValThr 490
QY 1430 CTCGACAAACACAAAC-----CGC 1447
Db LeuAlaProAspThrProGlyAlaAlaGlnThrThrGlyAlaGlyAlaProAlaTyrPro 510
QY 1448 CTGCAATGGCAGATCCCGCCACCCGA-----AAAACCTTACCGCAACGACCGGAGCC 1501
```

```
Db 511 LeuValTyrThrGluIleAlaAspGlyGlnValArgThrTyrSerValAspProAsnAla 530
QY 1502 AAACCTTTGGAAACCATCGCGCAAAATCTATCTCTGCTGCGCATCGAAGGTTTATTA 1561
Db 531 GlyPheTyrArgAsnLeuLeuThrGlyLeuCysLeuLeuLeuProIleAspGlnLeu 550
RESULT 38
Q88CA5_PSEPK
ID Q88CA5_PSEPK PRELIMINARY; PRT; 517 AA.
AC Q88CA5;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Phospholipase D family protein.
GN OrderedLocusNames=PP5276;
OS Pseudomonas putida (strain KT2440).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=160488;
[1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22433060; PubMed=12534463;
DOI=10.1046/j.1462-2920.2002.00366.x;
RA Nelson K.E., Weinel C., Paulsen I.T., Dodson R.J., Hilbert H.,
RA Martins dos Santos V.A.P., Fouts D.E., Gill S.R., Pop M., Holmes M.,
RA Brinkac L.M., Beanan M.J., DeBoy R.T., Daugherty S.C., Kolonay J.F.,
RA Madupu R., Nelson W.C., White O., Peterson J.D., Khouri H.M.,
RA Hance I., Chris Lee P., Holtzapfel E.K., Scanlan D., Tran K.,
RA Moazzez A., Uterback T.R., Rizzo M., Lee K., Kosack D., Moestl D.,
RA Wedler H., Lauber J., Stjepandic D., Hohnselt J., Straetz M., Heim S.,
RA Kiewitz C., Eisen J.A., Timmis K.N., Duesterhoeft A., Tuemmler B.,
RA Fraser C.M.;
RT "Complete genome sequence and comparative analysis of the
RT metabolically versatile Pseudomonas putida KT2440.";
RL Environ. Microbiol. 4:799-808(2002).
DR EMBL; AE016793; AAN70841.1; -; Genomic_DNA.
DR TIGR; PP5276; -.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR GO; GO:0008654; P:phospholipid biosynthesis; IEA.
DR InterPro; IPR001736; PLD.
DR Pfam; PF00614; PLDC; 2.
DR SMART; SM00155; PLDC; 2.
DR PROSITE; PS50035; PLD; 2.
KW Complete proteome.
SQ SEQUENCE 517 AA; 57574 MW; 107E07900AA35751 CRC64;
```

Alignment Scores:

Pred. No.:	5,79e-35	Length:	517
Score:	773.00	Matches:	199
Percent Similarity:	53.6%	Conservative:	83
Best Local Similarity:	37.8%	Mismatches:	210
Query Match:	27.1%	Indels:	34
DB:	2	Gaps:	15

US-10-665-990A-13 (1-1561) x Q88CA5_PSEPK (1-517)

```
QY 47 CGAGCGCTCATTCCTTTTATGCTCTCTC-----TGTTTCATGTTCTTCATGCTG 100
Db 5 ArgAlaLeuProLeuLeuValLeuLeuGlyValAlaGlyCysThrSerIleSer 24
QY 101 CCCCACTGGAGAAACGAGCGGAAACCGTCATTTCAATACTTCCAAACCTGCTCCTG 160
Db 25 AlaPro-----ArgGluThrSerGlnAlaLeuProAlaHisGluSerAlaPheGly 41
QY 161 GACAACATCTCGCAATTCGGCACACCCCTCATCAACACGGGCTATCCGACATCTACCTG 220
Db 42 ArgSerValLeu-----ArgGlnAlaAlaProTyrGlyArgSerGlyPheArgLeu 59
QY 221 CTCAGCAGCCCCACGAAAGCCCTTGGCGCGCGCCCTTATCGAATCTGCCGACAC 280
Db 60 LeuProAsnSerAsnGluAlaPheAlaArgAlaGluLeuIleArgAsnAlaGlnAla 79
```


Db	41	ProLaThrSerLeuAspArgProValSerHisAlaLeuSerAlaGluThrAlaThrPro	60
Qy	158	CTGGACAAACATCTGCAAAATCCGGCACACCCCTCATAACACGGGCTATCCGACATCTAC	217
Db	61	LeuArgAspAlaLeuAlaAlaProGluAlaHis---ProGlyGlnSerGlyPheArg	79
Qy	218	CTGCTCGACGACCCCGACGAAAGCCCTTCGCCGCCGCCGCCCTTATCGAATCTCCGCAA	277
Db	80	LeuLeuAlaAspGlyAlaThrAlaLeuGlnMetArgIleAlaLeuAlaArgAlaAlaThr	99
Qy	278	CACAGCCTCGAATTTGCATATCTACATTTTGGCGCAACGACATTTCCGGCAGGCTGCTTC	337
Db	100	LysThrLeuAspMetGlnTyrTyrIleAlaThrGluAspThrThrGlyLysLeuLeuLeu	119
Qy	338	AACCTCATGTCCTTGGCCGAGAACCGCGCTGGCGGTACGCTGCTCTTGGAGCACAAC	397
Db	120	GlyAlaAlaLeuTyrAlaAlaAspArgGlyValArgValaRgmMetLeuValaAspLeu	139
Qy	398	AACAGCGCGGGTTGGAGCATCTCTGCTCGCCCTCGACAGCCATCCCAATATCGAAGTG	457
Db	140	AsnPheHisAspIleAspArgValMetAlaAlaLeuAsnThrHisGlnAsnIleGluIle	159
Qy	458	CGCCTGTTCACCCCTTCGTCTACGGCAA-----TGGCGGCGACTCGGTAC	505
Db	160	ArgValPheAsnProPheGlyAlaSerGlnArgGlyMetMetGluArgThrAlaAsnPhe	179
Qy	506	CTGACCGCATTTCCCGCGCTCAAACCGCGCATGTACAAACAATCTTTTACCGCGCACAAC	565
Db	180	PheThrArgIleAspSerPheThrArgArgMetHisAsnLysAlaMetIleAlaAspAsn	199
Qy	566	CGCGCCACCATACTCGGCGGACGAATATCGCGACGCAATACTTCAAAGTCGGTAGGAC	625
Db	200	GlnLeuAlaIleValGlyGlyArgAsnLeuGlyAspGluTyrPheSerAlaSerProThr	219
Qy	626	ACCGTTTTCGCCGACCTGGACATCTTCGCCACCGCGCAGCTGCTCGGGCAAGTATCGCAC	685
Db	220	LeuGlnPheArgAspLeuAspValLeuAlaGlyProValThrSerAspIleSerLys	239
Qy	686	GACTTCCACCGCTACTGGCAAGCATTTCCGCCCACACGCGCCGCGCATCATCCGACG	745
Db	240	SerPheAspAspTyrTrpAlaSerAlaSerSerTyr---ProLeuArgValLeuAsnHis	258
Qy	746	GGCAACATCGGC---AAGGGTCTTCAAGCACTCGGTATACACGACGAAACATCCAGACAC	802
Db	259	GlnThrPheAspProLysAspLeuAspAlaMet-----ArgAspGluLeuArgAspHis	276
Qy	803	CGGCTCTCGCTACCGCGAAACCGTCGAA-----CAGTCG	838
Db	277	-----TrpArgLysAsnAlaAspProTyrAsnAlaLysProLeuAsnAlaThr	292
Qy	839	CCCCTCTACCAAAAATACAGCGGAGCGCATCGAC-----TGCGAGAGCGTCCAAACC	892
Db	293	ProLeuAlaArgGlnIleAlaArgAspGluLeuGluLeuValTrpAlaProAlaGluPhe	312
Qy	893	CGCCTGTATGACGACACCCCTGCAAAAAGGACTCGACCGCGACCCCGC-----	940
Db	313	LysVal-----AspAlaPro-----AspLysValAlaArgProThrGlyThr	326
Qy	941	-----AAACCGCGGATTGGCGGAGGCTGCAAGACGGCTCAACACGCCCGCAAAAAGC	994
Db	327	TyrValSerProProMetGln---ArgLeuAlaGluLeuThrArgGlyAlaGlnLysGlu	345
Qy	995	GTCTATCTGGTTTCACCCCTATTTTCGTCCCTACAAAATCCGCGACAGACGCTCGCAAAA	1054
Db	346	PheLeuAlaPheSerProTyrPheValProHisAspAlaGlyValLysIleLeuGlyAsp	365
Qy	1055	CTGGTCGAGCAGCATAGACGTTTACCGTCTCTGACCAACTCGCTACAGGGCGACGCGT	1114
Db	366	ThrValAlaArgGlyValArgValAlaIleValThrAsnSerLeuAlaAlaThrAspAla	385
Qy	1115	CGCGCGTCTCATTCGGCTACGTCAATACGGAACACCGTGTCTCAAGCGCGCATCAA	1174

Db 386 ValAlaValGlnSerGlyTyrAlaProTyrArgValProLeuLeuGlnArgGlyValGlu 405

QY 1175 CTCCTACGAGCTGCACCAACCATCGCGTCCCGCCACAAAGACAAAGCGCTGACCGCG 1234

Db 406 LeuTyrGluPheIys-----SerGlnProAspGlnGlnProAlaArgLeuPheGly 422

QY 1235 AGC---TCCGTGAACAGCCTGCATGCCAAACCTTCATTGTGGACGCCAAACGCATCTTC 1291

Db 423 SerArgSerArgAlaSerLeuHisAlaLysAlaTyrValIleAspArgGlnIleLeuVal 442

QY 1292 ATCGGCTCATTAACCTTCGACCCCGCTTCGCGACGGCTCAATACCGAAATGGGGCGTCGT 1351

Db 443 IleglySerLeuAsnLeuAspProArgSerAlaHisLeuAsnThrGluLeuAlaLeuVal 462

QY 1352 ATCGAAAGCCCCAAAATCGCAGAACAGATGGAGCGACCCCTCGCCGATACCACAC---CCC 1408

Db 463 IleHisSerProValLeuAlaGlnGlnAlaAlaIlePheAlaArgValThrGlnPro 482

QY 1409 GAATACCGCTACCGCGTTTACCCTTCGACAAACACAAAC-----CGCCTGCAA 1453

Db 483 AspGluSerTyrArgValThrLeuAlaLysArgThrAspGlyGlyProProAlaLeuGlu 502

QY 1454 TGG-----CACGATCCCGCCACCGGAAACCTACCCGACGACCGAACCGCAACTT 1507

Db 503 TrpThrGlyThrGluGlyGlyGlnThrAlaThrTyrHisValAspProHisAlaGlyLeu 522

QY 1508 TGGAAACGCATCGCGCAAAATCCTATCCCTCGCTGCCCATCGAAGTTTATTATTA 1561

Db 523 LeuArgAsnValMetThrGlyIlePheThrLeuLeuProValAspAspGlnLeu 540

RESULT 40

Q88BC2 PSESM

ID Q88BC2 PSESM PRELIMINARY; PRT; 542 AA.

AC Q88BC2;

DT 01-JUN-2003 (TrEMBLrel. 24, Created)

DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)

DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)

DE Phospholipase D family protein.

GN OrderedLocusNames=PSPT00095;

OS Pseudomonas syringae (pv. tomato).

OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;

OC Pseudomonadaceae; Pseudomonas.

EN NCBI_TaxID=323;

FN [1]

RC NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].

RC STRAIN=DC3000;

RX MEDLINE=22834015; PubMed=12928499; DOI=10.1073/pnas.1731982100;

RA Buell C.R., Joardar V., Lindeberg M., Selengut J., Paulsen I.T.,

RA Winn M.L., Dodson R.J., DeBoy R.T., Durkin A.S., Kolonay J.F.,

RA Madupu R., Daugherty S.C., Brinkac L.M., Beanan M.J., Haft D.H.,

RA Nelson W.C., Davidson T.M., Zafraac N., Zhou L., Liu J., Yuan Q.,

RA Khouri H.M., Fedorova N.B., Tran B., Russell D., Berry K.J.,

RA Uterback T.R., Van Aken S.E., Feldblyum T.V., D'Ascenzo M.,

RA Deng W.-L., Ramos A.R., Alfano J.R., Cartinhour S., Chatterjee A.K.,

RA Delaney T.P., Lazarowitz S.G., Martin G.B., Schneider D.J., Tang X.,

RA Bender C.L., White O., Fraser C.M., Collmer A.;

RT "The complete genome sequence of the Arabidopsis and tomato pathogen

PT Pseudomonas syringae pv. tomato DC3000.";

RL Proc. Natl. Acad. Sci. U.S.A. 100:10181-10186(2003).

DR EMBL; AE016853; AAO53649.1; -; Genomic_DNA.

DR TIGR; PSPT00095; -.

DR GO; GO:0003824; F:catalytic activity; IEA.

DR GO; GO:0008152; P:metabolism; IEA.

DR InterPro; IPR001736; PLD.

DR Pfam; PF00614; PLDC; 2.

DR SMART; SM00155; PLDC; 2.

DR DR PROSITE; PS00035; PLD; 2.

KW Complete proteome.

SQ SEQUENCE 542 AA; 60833 MW; 8B96B8F00C45F7B8 CRC64;

Alignment Scores: 9.73e-35 Length: 542

Pred. No.: 769.00 Matches: 200

Search completed: May 2, 2006, 05:24:25
Job time : 460 secs

Qy	701	TGGCAAGCATTCCGCCCAACAACGCCACGCGCATCATCCGACGGCGAAACATCGGCAAG	760
Db	240	TrpAsnSer-----ThrLeuSerLys	246
Qy	761	GGTCTTCACGACTCGGATAC-----AACGACGAAACATCCAGACACGCGCTCCTG	811
Db	247	ProileAspGlnPheMetTyrPheLeuProAspGlyGlnAspLeuAlaGluAlaArgGln	266
Qy	812	CGCTACCGCGAAACCGTCGAAACAGTCG-----CCCTCTACCAAAAAATA	856
Db	267	ArgLeuAspSerLeuGluGlnAlaHisGlnGlnHisLysAlaLeuTyrGluArgLeu	286
Qy	857	CAGACGGGACGCATCGAC-----TGGCAGAGC-----GTCCAA	889
Db	287	MetAlaTyrLysThrGlnProArgMetLysThrTyrLeuAsnGluLeuValTyrAlaHis	306
Qy	890	ACCGCGCTGATCAGCGACACCCCTGCAAAAGGACTCGACCGCGACCGCGCAAAACCGCGC	949
Db	307	AsnGlnAlaLeuTyrPAspAlaProThrLysValLeuAlaGlnGlyGluProAspProHis	326
Qy	950	-----ATTGCCGGGAGGCTGCAGAAGCGGCTCAAAACAGCCCGCAAAAAAAGCGTCTATCTG	1003


```
QY 122 GAAAGCGCTCATTTCAATCTCCAAACCTGTCCTCTCGACAAACATCTCGAAATCCGG 181
Db 46 GluSerArgHisPheAsnThrSerLysProValArgLeuAspAsnIleLeuGlnIleArg 65
QY 182 CACACCCCTCATAACACGGGCTATCCGACATCTACCTGCTCGACGACCCCGACGAGCC 241
Db 66 HisThrProHisThrAsnGlyLeuSerAspIleTyrLeuLeuAsnAspProHisGluAla 85
QY 242 CTTGCGCGCGCGCGCTTATCGAATCTGCGGAACACAGCCTCGATTTGCAATACTAC 301
Db 86 PheAlaAlaArgAlaAlaLeuIleGluSerAlaGluHisSerLeuAspLeuGlnTyrTyr 105
QY 302 ATTTGGCGAACAGCATTTCCGGCAGGCTGTGTTCAACCTCATGTACTTGGCCGAGAA 361
Db 106 IleTyrAsnAspIleSerGlyArgLeuLeuPheAsnLeuValTyrLeuAlaAlaGlu 125
QY 362 CCGCGGCTGCGGTAGGCTGTGTTGGACGACACACACGCGGGGTTGGACGATCTC 421
Db 126 ArgGlyValArgValArgLeuLeuLeuAspAsnAsnThrArgGlyLeuAspAspLeu 145
QY 422 CTGCTCGCCCTCGACAGCCATCCCAATATCGAAGTGGCGCTGTTCAACCCCTTCGTCTA 481
Db 146 LeuLeuAlaLeuAspSerHisProAsnIleGluValArgLeuPheAsnProPheValLeu 165
QY 482 CGCAATGCGCGCACTCGGCTACTGACCGACTTCCCGCGCTCAACCGCGCATGCAAC 541
Db 166 ArgLysTyrArgAlaLeuGlyTyrLeuThrAspPheProArgLeuAsnArgArgMethHis 185
QY 542 ACAAAATCTTTACCGCGACACCGCGCCACCATCTACCTCGGCGGACGCAATATCGGCAC 601
Db 186 AsnLysSerPheThrAlaAspAsnArgAlaThrIleLeuGlyGlyArgAsnIleGlyAsp 205
QY 602 GAATACTTTCAAAGTCGGTGGACACCGGTTTTCGCGCAGCTCGGACATCTCGCCACCGGC 661
Db 206 GluTyrPheLysValGlyGluAspThrValPheAlaAspLeuAspIleLeuAlaThrGly 225
QY 662 AGCGTGTGCGGGAAGTATCGCACGACTTCGACCGCTACTGGGGAAGCGCATTCGCGCCAC 721
Db 226 SerValValGlyGluValSerHisAspPheAspArgTyrTrpAlaSerHisSerAlaHis 245
QY 722 AACGCCACGCGCATCATCGCAGCGGCAACATCGCAAGGTCCTCAAGCACTCGGATAC 781
Db 246 AsnAlaThrArgIleIleArgSerGlyAsnIleGlyLysGlyLeuGlnAlaLeuGlyTyr 265
QY 782 AACGCAAAACATCCAGACACCGCTCTCTGCGCTACCGCGAACCGTTCGAAACAGTCGCC 841
Db 266 AsnAspGluThrSerArgHisAlaLeuLeuArgTyrArgGluThrValGluGlnSerPro 285
QY 842 CTCTACCAAAAATACAGACGGGACGATCGACTGGCAGAGCTCCAAACCGCGCTGATC 901
Db 286 LeuTyrGlnLysIleGlnThrGlyArgIleAspTyrGlnSerValGlnThrArgLeuIle 305
QY 902 AGCGACACCCCTGCAAAAGGACTTCGACCGCGACCGCGCAACCGCGATTCGCGGGAGG 961
Db 306 SerAspAspProAlaLysGlyLeuAspArgAspArgLysProProIleAlaGlyArg 325
QY 962 CTGCAAGACGCGCTCAAAACAGCCCGCAAAAACCGTCTATCTGGTTTCACCCCTATTTCGTC 1021
Db 326 LeuGlnAspAlaLeuLysGlnProGluLysSerValTyrLeuValSerProTyrPheVal 345
QY 1022 CTTCAAAAATCCGGCACAGACGACTGGGCAAAACCTGGTCGACGAGCGCATAGACGTTACC 1081
Db 346 ProThrLysSerGlyThrAspAlaLeuAlaLysLeuValGlnAspGlyIleAspValThr 365
QY 1082 GTCCTGACCACTCGCTACAGCGGACGAGCTTGGCGCGTCCATTTCGGGCTACGTCAAA 1141
Db 366 ValLeuThrAsnSerLeuGlnAlaThrAspValAlaAlaValHisSerGlyTyrValLys 385
QY 1142 TACCGAAAAACCGCTGTCAAAGCCGGCATCAAACTCTACGAGCTGCAACCAACCATGCC 1201
Db 386 TyrArgLysProLeuLeuLysAlaGlyIleLysLeuTyrGluLeuGlnProAsnHisAla 405
```

```
QY 1202 GTCCCCGCCACAAAAGAGAGGCGCTGACCGGAGCTCCGTAAACAGCGCTCGATGCCAAA 1261
Db 406 ValProAlaThrLysAspLysGlyLeuThrGlySerSerValThrSerLeuHisAlaLys 425
QY 1262 ACCTTTCATTGTGGAGCGCAACCGCATCTTATCGGCTCATTTCAACCTCGACCCCGTTCC 1321
Db 426 ThrPheIleValAspGlyLysArgIlePheIleGlySerPheAsnLeuAspProArgSer 445
QY 1322 GCACGGCTCAATACCGAAATCGGCGTCGTATCGAAAGCCCCCAAAATCGCAGAACAGATG 1381
Db 446 AlaArgLeuAsnThrGluMetGlyValIleGluSerProLysIleAlaGluGlnMet 465
QY 1382 GAGCGCACCTCGCGGATACCAACCCGAAATACCGCTACCGGTTACCTCGACAAAACAC 1441
Db 466 GluArgThrLeuAlaAspThrSerProGluTyrAlaTyrArgValThrLeuAspArgHis 485
QY 1442 AACCGCTGCAATGCGACGATCCGCGCACCCGCAAAACCTACCCGAAACGAAACCCGAAGCC 1501
Db 486 AsnArgLeuGlnTyrHisAspProAlaThrArgLysThrTyrProAsnGluProGluAla 505
QY 1502 AAACCTTTGGAAACGATCGCGCAAAAATCTATCCCTGCTGCCCATCGAAGTTTATTA 1561
Db 506 LysLeuTyrLysArgIleAlaAlaLysIleLeuSerLeuLeuProIleGluSerLeuLeu 525
RESULT 2
BB1083
cardiolipin synthetase family protein NMB1434 [imported] - Neisseria meningitidis (strai
C:Species: Neisseria meningitidis
C:Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 09-Jul-2004
C:Accession: BB1083
R:Tetelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, J.A.;
Hickey, E.K.; Haft, D.H.; Salzberg, S.B.; White, O.; Fleischmann, R.D.; Dougherty, B.A.;
ri, H.; Qin, H.; Vamathevan, J.; Gill, J.; Scarlato, V.; Masigiani, V.; Pizza, M.
Science 287, 1809-1815, 2000
A:Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.; Ve
A:Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MC58.
A:Reference number: AB1000; MUID:20175755; PMID:10710307
A:Accession: BB1083
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-508 <TET>
A:Cross-references: UNIPROT:Q9JYU0; UNIPARC:UPI00000C46E9; GB:AE002494; GB:AE002098; NID
A:Experimental source: serogroup B, strain MC58
C:Genetics:
A:Gene: NMB1434
C:Superfamily: cardiolipin synthase
Alignment Scores:
Pred. No.: 1,93e-143 Length: 508
Score: 2573.00 Matches: 497
Percent Similarity: 98.8% Conservative: 5
Best Local Similarity: 97.8% Mismatches: 6
Query Match: 90.2% Indels: 0
DB: 2 Gaps: 0
US-10-665-990A-13 (1-1561) x BB1083 (1-508)
QY 38 ATGAAAAACGACGAGCCTCATTTCCCTTTATGCGCTCCTCTCTGTTCATGTTCTTCATCG 97
Db 1 MetLysThrArgSerLeuIleSerLeuLeuCysLeuLeuLeuCysSerCysSerSerTrp 20
QY 98 TTGCCCCCAGCTCGAGAACGAGGAGAGCGGTCTATTTCAATACTTCCAAACCTGTCCTCC 157
Db 21 LeuProProLeuGluGluArgThrGluSerArgHisPheAsnThrSerLysProValArg 40
QY 158 CTGGACAAACATCCTCAAAATCCGCGACACCCCTCATAAACACGCGGTATCCGACATCTAC 217
Db 41 LeuAspAsnIleLeuGlnIleArgHisThrProHisThrAsnGlyLeuSerAspIleTyr 60
QY 218 CTGCTCGACGACCCCGACGAGCCCTTCCGCGCGCGCGCCCTTATCGAATCTGCCGAA 277
Db 61 LeuLeuAsnAspProHisGluAlaPheAlaAlaArgAlaAlaLeuIleGluSerAlaGlu 80
```

```
QY 278 CACAGCCTCGATTTCGAATACATATTGGCGCAACAGCATTTCCGGCAGGCTGCTGTC 337
Db 81 HisSerLeuAspLeuGlnTyrTyrIleTrpArgAsnAspIleSerGlyArgLeuLeuPhe 100
QY 338 AACCTCATGTACCTTGGCGCAGACGGCGGTGGCGTACCGCTGCTTTGGACGACAAAC 397
Db 101 AsnLeuValTyrLeuAlaAlaGluArgGlyValArgValArgLeuLeuLeuAspAsn 120
QY 398 AACAGCGCGGGTTGGACGATCTCTCTGCTCGCCCTCGACAGCCATCCCAATATCGAAGTG 457
Db 121 AsnThrArgGlyLeuAspAspLeuLeuAlaLeuAspSerHisProAsnIleGluVal 140
QY 458 CGCTGTTCAACCCCTTCTGCTCTACGAAATGGCGGCACCTCGCTACCTGACCGACTTC 517
Db 141 ArgLeuPheAsnProPheValLeuArgLysTrpArgAlaLeuGlyTyrLeuThrAspPhe 160
QY 518 CCCCGCCTCAACCGCGCATGCACACAAATCTTTACCGCGGACACCGCGCCACCATTA 577
Db 161 ProArgLeuAsnArgMetHisAsnLysSerPheThrAlaAspAsnArgAlaThrIle 180
QY 578 CTCGGCGGACGCAATATCGCGACGAATACTTTCAAAGTCGGTGAGGACACCGCTTTTCGCC 637
Db 181 LeuGlyGlyArgAsnIleGlyAspGluTyrPheLysValGlyGluAspThrValPheAla 200
QY 638 GACCTGGACATCTCTGCCACCGCAGCGCTCGTCGGCGAAGTATCGCAGACTTCGACCGC 697
Db 201 AspLeuAspIleLeuAlaThrGlySerValValGlyGluValSerHisAspPheAspArg 220
QY 698 TACTGGGCAAGCCATTCCGCCACAAACCGCAGGCGCATCATCCGAGGGGGAACATCGGC 757
Db 221 TyrTrpAlaSerHisSerAlaHisAsnAlaThrArgIleIleArgSerGlyAspIleGly 240
QY 758 AAGCGTCTTCAAGCACTCGGATACAAACGACCAAAATCCACACACGCGCTCTCGGTAC 817
Db 241 LysGlyLeuGlnAlaLeuGlyTyrAsnAspGluThrSerArgHisAlaLeuLeuArgTyr 260
QY 818 CGCGAAACCGTCGAACAGTCGCCCTCTTACCAAAAAATACAGACGGGACGCATCGACTGG 877
Db 261 ArgGluThrValGluGlnSerProLeuTyrGlnLysIleGlnThrGlyCysIleAspTrp 280
QY 878 CAGAGCTCCAAACCCGCTGATCAGGACACCCCTCGAAAGACTCGACCGGACCGC 937
Db 281 GlnSerValArgThrArgLeuIleSerAspAspProAlaLysGlyLeuAspArgAspArg 300
QY 938 CGCAAAACCGCGATTGCGGGAGGCTGCAAGACGCGCTCAAAACGCGCGGAAAAAGGTC 997
Db 301 ArgLysProProIleAlaGlyArgLeuGlnAspAlaLeuLysGlnProGluLysSerVal 320
QY 998 TATCTGTTTCAACCTATTTCCCTTACAAAATCCGGCACAGACGCACTGGCAAAACTG 1057
Db 321 TyrLeuValSerProTyrPheValProThrLysSerGlyThrAspAlaLeuAlaLysLeu 340
QY 1058 GTGCAGACGGCATAGAGTTTACCGTCTGACCAACTCGCTACAGGGCAGCGAGTTGCC 1117
Db 341 ValGlnAspGlyIleAspValThrValLeuThrAsnSerLeuGlnAlaThrAspValAla 360
QY 1118 GCCGTCCATTCCGGCTACGTCAAATACGAAACCGGTGCTCAAGCGCGCATCAAACTC 1177
Db 361 AlaValHisSerGlyTyrValLysTyrArgLysProLeuLeuLysAlaGlyIleLysLeu 380
QY 1178 TAGAGCTGCNAACCAACCTACCGTCCCGCCCAAAAAGACAAAGCCCTGACCGGAGC 1237
Db 381 TyrGluLeuGlnProAsnHisAlaValProAlaThrLysAspLysGlyLeuThrGlySer 400
QY 1238 TCGGTACCGAGCTGCTGCAAAACCTTCAATTGTCAGCGGCAAAACCGCATCTTCATCGGC 1297
Db 401 SerValThrSerLeuHisAlaLysThrPheIleValAspGlyLysArgIlePheIleGly 420
QY 1298 TCATTCAACCTCGACCCCGTTCCGACAGCGCTCAATACCGAAATGGCGCTCGTCATCGAA 1357
Db 421 SerPheAsnLeuAspProArgSerAlaArgLeuAsnThrGluMetGlyValValIleGlu 440
QY 1358 AGCCCCAAAATCGCAGAACAGATGGAGCGCACCCCTCGCCGATACCAACCCGGAATAGCC 1417
```

```
Db 441 SerProLysIleAlaGluGlnMetGluArgThrLeuAlaAspThrProAlaTyrAla 460
QY 1418 TACCGCGTTACCTTCGACAAACACACCGCCTGCAATGGCAGATCCCGCCACCGAAAA 1477
Db 461 TyrArgValThrLeuAspArgHisAsnArgLeuGlnTrpHisAspProAlaThrArgLys 480
QY 1478 ACCTACCGCGAACGAAACCGAAGCCAAACTTTTGGAAACGATCGCGCAAAAATCCTATCC 1537
Db 481 ThrTyrProLanGluProGluAlaLysLeuTrpLysArgIleAlaAlaLysIleLeuSer 500
QY 1538 CTGCTGCCCATCGAAGGTTTATTA 1561
Db 501 LeuLeuProIleGluGlyLeuLeu 508
RESULT 3
C64847
YmdC protein - Escherichia coli (strain K-12)
C:Species: Escherichia coli
C:Date: 12-Sep-1997 #sequence_revision 17-Sep-1997 #text_change 09-Jul-2004
C:Accession: C64847
R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.;
.A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A:Title: The complete genome sequence of Escherichia coli K-12.
A:Reference number: A64720; MUID:97426617; PMID:9278503
A:Accession: C64847
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-493 <BLAT>
A:Cross-references: UNIPROT:P75919; UNIPARC:UPI000013B993; GB:AE0000206; GB:U000096; NID
A:Experimental source: strain K-12, substrain MG1655
C:Genetics:
A:Gene: ymdC
C:Superfamily: cardiolipin synthase
Alignment Scores:
Pred. No.: 3,32e-57 Length: 493
Score: 1102.50 Matches: 221
Percent Similarity: 65.5% Conservative: 83
Best Local Similarity: 47.6% Mismatches: 143
Query Match: 38.7% Indels: 17
DB: 2 Gaps: 5
US-10-665-990A-13 (1-1561) x C64847 (1-493)
QY 200 GGGCTATCCGACATCTACTGCTCGACGACCCCGACGACCTTGGCGCGCGCGCC 259
Db 37 GlyGlnCysGlyLeuPheProLeuGluLysSerLeuAspAlaAlaArgTyrArg 56
QY 260 CTTATCGAATCTCGGACACAGCCTCGATTTCGAATCTACATTTTGGCGCAACGACATT 319
Db 57 LeuAlaGluMetAlaGluHisThrLeuAspValGlnTyrTyrIleTrpGlnAspAspMet 76
QY 320 TCCGGCAGGCTGCTGTTTCAACCTCATGTACTTCCCGCAGAACGCGCGCTGCGGTACGC 379
Db 77 SerGlyArgLeuLeuPheSerAlaLeuAlaAlaLysArgGlyValArgValArg 96
QY 380 CTGCTGTTGACGACAAACACGCGCGGTTTGACGATCTCTCTGCTCGCCCTCGACAGC 439
Db 97 LeuLeuLeuAspAsnAsnThrProGlyLeuAspAspIleLeuArgLeuLeuAspSer 116
QY 440 CATCCCAATATCCAAGTGGCGCTTCAACCCCTTCTGCTCTACGCAATGGCGCGCATC 499
Db 117 HisProArgIleGluValArgLeuPheAsnProPheSerPheArgLeuLeuArgProLeu 136
QY 500 GGTCTACTCGACCATTTCCCGCCCTCAACCGCGCATGCACAAATCTTTTACCGCC 559
Db 137 GlyTyrIleThrAspPheSerArgLeuAsnArgArgMetHisAsnLysSerPheThrVal 156
QY 560 GACAACCGCCCACTACTCTCGGCGGACGCAATATCGGCGACGCAATATCTTCAAAGTCGT 619
Db 157 AspGlyValValThrLeuValGlyGlyArgAsnIleGlyAspAlaTyrPheGlyAlaGly 176
```

```
QY 620 GAGGACACGGTTTCCGCGACCTGACATCTCTCGCACCGGAGCGTGTCTGGCGAAGTA 679
Db 177 GluGluProLeuPheSerAspLeuAspValMetAlaIleGlyProValValGluAspVal 196
QY 680 TCGCAGCACTTCGACCGTACTGCGCAAGCCATTCCGCCCAACACCGCGCATCATC 739
Db 197 AlaAspAspPheAlaArgTyrTrpTyrCysLysSerValSerProLeuGlnGlnValLeu 216
QY 740 CGCAGCGGCAACATCGCAAGGGT-----CTTCAAGCATCTCGGA 778
Db 217 -----AspValProGluGlyGluMetAlaAspArgIleGluLeuProAlaSerTrp 233
QY 779 TACAACGAGCAACATCCAGACACCGCTCTCGCTACCGCGAAGCCGTCGAACAGTCG 838
Db 234 HisAsnAspAlaMetThrHis-----ArgTyrLeuArgLysMetGluSerSer 249
QY 839 CCCCTCTACCAAAATACAGACGGGACGATCGACTGGCAGCGTCCAAACCGCGCTG 898
Db 250 ProPheIleAsnHisLeuValAspGlyThrLeuProLeuIleTrpAlaLysThrArgLeu 269
QY 899 ATCAGCGACACCCCTGCAAAAGGACTCGACCGCGACCGCGCAACCCCGGATTCCCGG 958
Db 270 LeuSerAspProAlaLysGlyGluGlyLysAlaLysArgHisSerLeuLeuProGln 289
QY 959 AGGCTGCAAGCGCTCAACAGCGCGGAAAGCGTCTATCTGGTTTACCCTATTTC 1018
Db 290 ArgLeuPheAspIleMetGlySerProSerGluArgIleAspIleSerSerTyrPhe 309
QY 1019 GTCCCTTACAAAATCCGGCACACGACCTGCGCAAACTGGTCAGGAGCGGATAGCGTT 1078
Db 310 ValProThrArgAlaGlyValAlaGlnLeuLeuArgMetValArgLysGlyValLysIle 329
QY 1079 ACCGTCCTGACCAACTCGCTACAGCGCGACCGAGTGGCGCGTCCATTCGGCTACGTC 1138
Db 330 AlaIleLeuThrAsnSerLeuAlaAsnAspValAlaValAlaIleAlaGlyTyrAla 349
QY 1139 AAATACCGAAACCGCTGTCAAGCCGCGCATCAACTCTACAGCTGCAACCCCAACCAT 1198
Db 350 ArgTrpArgLysLeuLeuArgTyrGlyValGluLeuTyrGluLeuLysProThrArg 369
QY 1199 GCGCTCCCGGCACACA---AAAGACAAAGCGCTGACCGGAGCTCCGTAACAGCGCTGCAT 1255
Db 370 GluGlnSerSerThrLeuHisAspArgGlyIleThrGlyAsnSerGlyAlaSerLeuHis 389
QY 1256 GCACAAACCTTCATTGTGACGGCAACGCATCTTCATCGGCTCATCAACTCGACGCC 1315
Db 390 AlaLysThrPheSerIleAspGlyLysThrValPheIleGlySerPheAsnPheAspPro 409
QY 1316 CGTTCCGACGGCTCAATACCAATGGCGTGTCTCATCGAAGCCCAAAATCGCAGAA 1375
Db 410 ArgSerThrLeuLeuAsnThrGluMetGlyPheValIleGluSerGluThrLeuAlaGln 429
QY 1376 CAGATGGAGCGCACCTCCCGCATACACACCGCAATACGCTACCGGTACCTCGAC 1435
Db 430 LeuIleAspLysArgPheIleGlnSerGlnTyrAspAlaAlaTrpGlnLeuArgLeuAsp 449
QY 1436 AAACACAAACCGCTGCAATGGACGATCCCGCGACCCGCAAA-----ACCTACCGCAAC 1489
Db 450 ArgTrpGlyArgIleAsnTrpValAspArgHisAlaLysLysGluIleLeuLysLys 469
QY 1490 GAACCGGACCAAACTTGGAAACGCATCGCGCGCAAAATCCTATCCTGCTGCGCCATC 1549
Db 470 GluProAlaThrSerPheTrpLysArgValMetValArgLeuAlaSerIleLeuProVal 489
QY 1550 GAAGGTTTATTA 1561
Db 490 GluTrpLeuLeu 493
RESULT 4
H90806
probable synthase [imported] - Escherichia coli (strain O157:H7, substrain RIMD 0509952)
C.Species: Escherichia coli
```

```
C.Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 09-Jul-2004
C.Accession: H90806
R.Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A.Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and geno
A.Reference number: A99629; MUID:21156231; PMID:11258796
C.Accession: H90806
A.Status: preliminary
A.Molecule type: DNA
A.Residues: 1-493 <HAY>
A.Cross-references: UNIPROT:Q8X917; UNIPARC:UPI00000D07FD; GB:BA000007; PIDN:BA034847.1;
A.Experimental source: strain O157:H7, substrain RIMD 0509952
C.Genetics:
A.Gene: ECs1424
C.Superfamily: cardiolipin synthase
```

```
Alignment Scores:
Pred. No.: 4,35e-57 Length: 493
Score: 1100.50 Matches: 220
Percent Similarity: 65.5% Conservative: 84
Best Local Similarity: 47.4% Mismatches: 143
Query Match: 38.6% Indels: 17
DB: 2 Gaps: 5
```

US-10-665-990A-13 (1-1561) x H90806 (1-493)

```
QY 200 GGGTATCCGACATCTACTCTCTGACAGACCCCGCAAGCGCTTTCGCGCGCGCCGCC 259
Db 37 GlyGlnCysGlyLeuPheProLeuGluLysSerLeuAspAlaPheAlaAlaArgTyrArg 56
QY 260 CTTATCGAATCTCGCAACACAGCTCGATTGCAATACTACATTGGCGCAACACCAT 319
Db 57 LeuAlaGluMetSerGluHisThrLeuAspValGlnTyrTrpIleTrpGlnAspMet 76
QY 320 TCCGCGAGCGTGTCTTCAACTCATCTGTTCGCGAGAACGCGCGTGTGCGCTAGCG 379
Db 77 SerGlyArgLeuLeuPheSerAlaLeuLeuAlaAlaLysArgGlyValArgValArg 96
QY 380 CTGCTGTTGGACGACAAACACACGCGGGTTGACGATCTCTGCTCGCCTCGACAGC 439
Db 97 LeuLeuLeuAspAspAsnAsnThrProGlyLeuAspAspIleLeuArgLeuLeuAspSer 116
QY 440 CATCCCAATATCGAAGTGGCGCTGTTCACCCCTTCGCTCTACGCAAAATGGCGCGCATC 499
Db 117 HisProArgIleGluValArgLeuPheAsnProPheSerPheArgLeuLeuArgProLeu 136
QY 500 GGCTACCTGACGACTTCCCGCGCTCAACCGCGCATGCACAAATCTTTTACCGCC 559
Db 137 GlyTyrIleThrAspPheSerArgLeuAsnArgArgMetHisAsnLysSerPheThrVal 156
QY 550 GACAAACCGCGCCACCATCTCGCGGACGCAATATCGCGGAGCAATATCTTCAAGTCGGT 619
Db 157 AspGlyValValThrLeuValGlyGlyArgAsnIleGlyAspAlaTyrPheGlyAlaGly 176
QY 620 GAGGACACGGTTTTCGCGACCTGGACATCTCGCACCGCGGAGCGTGTTCGCGCAAGTA 679
Db 177 GluGluProLeuPheSerAspLeuAspValMetAlaIleGlyProValValGluAspVal 196
QY 680 TCGCAGCACTTCGACCGTACTTGGGCAAGCCATTCCGCCCAACACCGCGCATCATC 739
Db 197 AlaAspAspPheAlaArgTyrTrpTyrCysLysSerValSerProLeuGlnGlnValLeu 216
QY 740 CGCAGCGCAACATCGCAAGGGT-----CTTCAAGCATCTCGGA 778
Db 217 -----AspValProGluGlyGluMetAlaAspArgIleGluLeuProAlaSerTrp 233
QY 779 TACAACGAGCAACATCCAGACACCGCTCTCGCTACCGCGAAGCCGTCGAACAGTCG 838
Db 234 HisAsnAspAlaMetThrHis-----ArgTyrLeuArgLysMetGluSerSer 249
QY 839 CCCCTCTACCAAAATACAGACGGGAGCGATCGACTGGCAGCGTCCAAACCGCGCTG 898
Db 250 ProPheIleAsnHisLeuValAspGlyThrLeuProLeuIleTrpAlaLysThrArgLeu 269
QY 899 ATCAGCGACACCCCTGCAAAAGGACTCGACCGCGACCGCGCAACCCCGGATTCCCGG 958
Db 270 LeuSerAspProAlaLysGlyGluGlyLysAlaLysArgHisSerLeuLeuProGln 289
QY 959 AGGCTGCAAGCGCTCAACAGCGCGGAAAGCGTCTATCTGGTTTACCCTATTTC 1018
Db 290 ArgLeuPheAspIleMetGlySerProSerGluArgIleAspIleSerSerTyrPhe 309
QY 1019 GTCCCTTACAAAATCCGGCACACGACCTGCGCAAACTGGTCAGGAGCGGATAGCGTT 1078
Db 310 ValProThrArgAlaGlyValAlaGlnLeuLeuArgMetValArgLysGlyValLysIle 329
QY 1079 ACCGTCCTGACCAACTCGCTACAGCGCGACCGAGTGGCGCGTCCATTCGGCTACGTC 1138
Db 330 AlaIleLeuThrAsnSerLeuAlaAsnAspValAlaValAlaIleAlaGlyTyrAla 349
QY 1139 AAATACCGAAACCGCTGTCAAGCCGCGCATCAACTCTACAGCTGCAACCCCAACCAT 1198
Db 350 ArgTrpArgLysLeuLeuArgTyrGlyValGluLeuTyrGluLeuLysProThrArg 369
QY 1199 GCGCTCCCGGCACACA---AAAGACAAAGCGCTGACCGGAGCTCCGTAACAGCGCTGCAT 1255
Db 370 GluGlnSerSerThrLeuHisAspArgGlyIleThrGlyAsnSerGlyAlaSerLeuHis 389
QY 1256 GCACAAACCTTCATTGTGACGGCAACGCATCTTCATCGGCTCATCAACTCGACGCC 1315
Db 390 AlaLysThrPheSerIleAspGlyLysThrValPheIleGlySerPheAsnPheAspPro 409
QY 1316 CGTTCCGACGGCTCAATACCAATGGCGTGTCTCATCGAAGCCCAAAATCGCAGAA 1375
Db 410 ArgSerThrLeuLeuAsnThrGluMetGlyPheValIleGluSerGluThrLeuAlaGln 429
QY 1376 CAGATGGAGCGCACCTCCCGCATACACACCGCAATACGCTACCGGTACCTCGAC 1435
Db 430 LeuIleAspLysArgPheIleGlnSerGlnTyrAspAlaAlaTrpGlnLeuArgLeuAsp 449
QY 1436 AAACACAAACCGCTGCAATGGACGATCCCGCGACCCGCAAA-----ACCTACCGCAAC 1489
Db 450 ArgTrpGlyArgIleAsnTrpValAspArgHisAlaLysLysGluIleLeuLysLys 469
QY 1490 GAACCGGACCAAACTTGGAAACGCATCGCGCGCAAAATCCTATCCTGCTGCGCCATC 1549
Db 470 GluProAlaThrSerPheTrpLysArgValMetValArgLeuAlaSerIleLeuProVal 489
QY 1550 GAAGGTTTATTA 1561
Db 490 GluTrpLeuLeu 493
```

Db	250	ProPheHleAsnHisLeuValAspGlyThrLeuProLeuIleTrpAlaLysThrArgLeu	269
QY	899	ATCAGCGACACCCCTGCAAAAGGATCTGACCGCGACCGCGCAAAACCGCGATTCCCGGG	958
Db	270	LeuSerAspProAlaLysGlyGluGlyLysAlaLysArgHisSerLeuLeuProGln	289
QY	959	AGGCTGCAAGCGCGCTCAACAGCGCGCAAAAGCGTCTATCTGTTTCAACCCATTTC	1018
Db	290	ArgLeuPheAspIleMetGlySerProSerGluArgIleAspIleSerSerTyrPhe	309
QY	1019	GTCCCTTACAAATCCGCGCACAGCCACTGCGCAAACTGGTCGAGGAGCGGATAGACGTT	1078
Db	310	ValProThrArgAlaGlyValAlaGlnLeuLeuArgMetValArgLysGlyValLysIle	329
QY	1079	ACCGTCTGACCACTCGCTACAGCGGACCGAGCTTGGCGCGTCCATCTCCGGTCTACGTC	1138
Db	330	AlaIleLeuThrAsnSerLeuAlaAlaAsnAspValAlaValHisAlaGlyTyrAla	349
QY	1139	AAATACCGAAACCGCTGCTCAAGCCGCGCATCAAACTCTACGAGCTGCAACCCCAACCAT	1198
Db	350	ArgTrpArgLysLysLeuLeuArgTyrGlyValGluLeuTyrGluLeuLysProThrArg	369
QY	1199	GCGGTCGCCGCACA--AAAGACAAAGCGCTGACCGCGAGCTCCGTAAACGAGCTGCAT	1255
Db	370	GluGlnSerSerThrLeuHisAspArgGlyIleThrGlyAsnSerGlyAlaSerLeuHis	389
QY	1256	GCCAAACCTTCATTGTGAGCGGCAACGCATCTTCATCGGCTCATTCACACTCGACCCC	1315
Db	390	AlaLysThrPheSerIleAspGlyLysThrValPheIleGlySerPheAsnPheAspPro	409
QY	1316	CGTTCCGCGACGGCTCAATACCGAAATGGCGTGTGATCGAAAGCGCCCAAAATCGCGAA	1375
Db	410	ArgSerThrLeuLeuAsnThrGluMetGlyPheValIleGluSerGluThrLeuAlaGln	429
QY	1376	CAGATGGAGCCACCTCGCGCATACACACCGCAATACGCTACCGGTTACCTCGAC	1435
Db	430	LeuIleAspLysArgPheIleGlnSerGlnTyrAspAlaAlaTrpGlnLeuArgLeuAsp	449
QY	1436	AAACACAAACCGCTGCAATGCGACGATCCCGCCACCGCAAA--ACCTACCGCAAC	1489
Db	450	ArgTrpGlyArgIleAsnTrpValAspArgHisAlaLysLysGluIleValLeuLysLys	469
QY	1490	GAACCGGAGCCAAACTTTGAAACGCGATCCCGCGCAAAATCTTATCCCTGCTCCCATC	1549
Db	470	GluProAlaThrSerPheTrpLysArgValMetValArgLeuAlaSerIleLeuProVal	489
QY	1550	GAAGGTTTATTA 1561	
Db	490	GluTrpLeuLeu 493	
RESULT 5			
D8566			
Probable synthase ymdC [imported] - Escherichia coli (strain O157:H7, substrain EDL933)			
C;Species: Escherichia coli			
C;Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004			
C;Accession: D8566			
R;Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew			
iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimantanta, E.; Potamouisis, K.; Apodaca,			
Nature 409, 529-533, 2001			
A;Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.			
A;Reference number: A85480; MUID:21074935; PMID:11206551			
A;Status: preliminary			
A;Molecule type: DNA			
A;Residues: 1-493 <STO>			
A;Cross-references: UNIPROT:Q8X9I7; UNIPARC:UPI0000165766; GB:AB005174; NID:gl2514578; F			
A;Experimental source: strain O157:H7, substrain EDL933			
C;Genetics:			
A;Gene: ymdC			
C;Superfamily: cardiolipin synthase			
Alignment Scores:			
Pred. No.: 1.12e-56 Length: 493			

Score: 1093.50 Matches: 219
Percent Similarity: 65.3% Conservative: 84
Best Local Similarity: 47.2% Mismatches: 144
Query Match: 38.3% Indels: 17
DB: 2 Gaps: 5

US-10-665-990A-13 (1-1561) x D85666 (1-493)

QY	200	GGGCTATCCGACATCTACCTCTCGACGACCCCGACAGCCCTTGCCTGCGCGCGCCGCC	259
Db	37	GlyGlnCysGlyLeuPheProLeuGluLysSerLeuAlaPheAlaAlaArgTyrArg	56
QY	260	CTTATCGAATCTGCGCAACACAGCTCGATTGCAATACTACTATTGGCGCAACACAT	319
Db	57	LeuAlaGluMetSerGluHisThrLeuAspValGlnTyrTyrIleTrpGlnAspMet	76
QY	320	TCCGGCAGGCTGTGTTCAACCTCATCTTCCCGGAGAACCGCGGTGGGTACGC	379
Db	77	SerGlyArgLeuLeuPheSerAlaLeuAlaAlaLysArgGlyValArgValArg	96
QY	380	CTGCTGTGGACGACAAACACCGCGGGTGGACGATCTCTCTGCTCGCCCTCGACGC	439
Db	97	LeuLeuLeuAspAspAsnAsnThrProGlyLeuAspIleLeuArgLeuLeuAspSer	116
QY	440	CATCCCAATATCGAAGTGGCGCTCTTCAACCCCTTCGCTCTACGCAAAATGGCGCACTC	499
Db	117	HisProArgIleGluValArgLeuPheAsnProPheSerPheArgLeuLeuArgProLeu	136
QY	500	GGCTACTGACCGATTCTCCCGCTCAACCGCGCATGCAACAACAATCTTTACCGCC	559
Db	137	GlyTyrIleThrAspPheSerArgLeuAsnArgArgMethHisAsnLysSerPheThrVal	156
QY	560	GACAAACCGCCCATCTACTCGCGGACGCAATATCGGACGCAATATCTTCAAAGTCGT	619
Db	157	AspGlyValValThrLeuValGlyArgAsnIleGlyAspAlaTyrPheGlyAlaGly	176
QY	620	GAGGACACCGTTTTCGCGCATCTCTCGCCACCGCGACGCTCGTGGCGAACTA	679
Db	177	GluGluProLeuPheSerAspLeuAspValMetAlaIleGlyProValValGluAspVal	196
QY	680	TGCGACGACTTGCACCGCTACTGGCAAGCATTCGCCCCCAACACGCGCATCATC	739
Db	197	AlaAspAspPheAlaArgTyrTrpTyrCysLysSerValSerProLeuGlnValLeu	216
QY	740	CGCAGCGCAACATCGCAAGGT-----CCTCAAGCACTCGA 778	
Db	217	-----AspValProGluGlyGluMetAlaAspArgIleGluLeuProAlaSerTrp	233
QY	779	TACAACGACGAAACATCCAGACACGCGCTCTCGCTACCGCAAAACCGTGAACAGTCG	838
Db	234	HisAsnAspAlaMetThrHis-----ArgTyrLeuArgLysMetGluSerSer	249
QY	839	CCCCTTACCAAAATACAGAGCGGACGATCGACTGGCAGAGCGTCCAAACCGCGCTG	898
Db	250	ProPheIleAsnHisLeuValAspGlyThrLeuProLeuIleTrpAlaLysThrArgLeu	269
QY	899	ATCAGCGACACCTCGCAAGGACTCGACCGGACCGCGCGCAACCGCGATTGCCGGG	958
Db	270	LeuSerAspAspProAlaLysGlyGlyGlyLysAlaLysArgHisSerLeuLeuProGln	289
QY	959	AGGCTGCAAGCGCTCAAAACAGCCCGAAAGCGCTCTATCTGTTTCAACCCATTTC	1018
Db	290	ArgLeuPheAspIleMetGlySerProSerGluArgIleAspIleSerSerTyrPhe	309
QY	1019	GTCCCTTACAAATCCGCGCACAGCCACTGCGCAAACTGGTCGAGGAGCGGATAGAGTT	1078
Db	310	ValProThrArgAlaGlyValAlaGlnLeuLeuArgMetValArgLys**ValLysIle	329
QY	1079	ACCGTCTGACCACTCGCTACAGGCGGACCGAGTTCGCCCGCTCCATCTCGGCTAGTC	1138
Db	330	AlaIleLeuThrAsnSerLeuAlaAlaAsnAspValAlaValHisAlaGlyTyrAla	349
QY	1139	AAATACCGAAACCGCTGCTCAAGCGCGCATCAAACTCTACGAGCTGCAACCCCAACCAT	1198

Db 350 ArgTrpArgLysLeuAraGlyGlyValGluLeuTyrGluLeuLysProThrArg 369
Qy 1199 GCCGTCCCGGCACA---AAAGACAAAGGGCTGACCGGACGCTCCGTAAACCAACGCTGCAT 1255
Db 370 GluGlnSerThrLeuHisAspArgGlyIleThrGlyAsnSerGlyAlaSerLeuHis 389
Qy 1256 GCGAAAACCTTCATTGTGGAGCGGAACGATCTTCATCGGCTCATTAACCTCGACCCC 1315
Db 390 AlalysThrPheSerIleAspGlyLysThrValPheIleGlySerPheAsnPheAspPro 409
Qy 1316 CGTTCCGACGCGCTCAATACCAGAAATGGCGTGTCATCGAAAGCCCCAACATCGCAGAA 1375
Db 410 ArgSerThrLeuLeuAsnThrGluMetGlyPheValIleGluSerGluThrLeuAlaGln 429
Qy 1376 CAGATGAGCGCACCTCGCGGATACACACACCGCAATACCGCTACCGCGTTACCTCGAC 1435
Db 430 LeuileAspLysArgPheIleGlnSerGlnTyrAspAlaIleTrpGlnLeuAsgLeuAsp 449
Qy 1436 AAACACAAACCGCTGCAATGGCAGCATCCGCCACCCGAAAA-----ACCTACCCGAAC 1489
Db 450 ArgTrpGlyArgIleAsnTrpValAspArgHisAlaLysLysGluLeuValLeuLysLys 469
Qy 1490 GAACCCGAAGCAAACCTTGGAAACGATCGCGCAAAATCTATCCCTGCTGCCCATC 1549
Db 470 GluProAlaThrSerPheTrpLysArgValMetValArgLeuAlaSerIleLeuProVal 489
Qy 1550 GAAGGTTTATA 1561
Db 490 GluTrpLeuLeu 493

RESULT 6
AE0636
conserved hypothetical protein STY1185 [imported] - Salmonella enterica subsp. enterica
C:Species: Salmonella enterica subsp. enterica serovar Typhi
A:Note: This species has also been called Salmonella typhi
C>Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 09-May-2004
C:Accession: AE0636
R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,
Th., T.; Connor, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,
S.S.; Moulé, S.; O'Gaora, P.
Nature 413, 848-852, 2001
A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;
A:Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serov
A:Reference number: AB0502; UID:2153494; PMID:11677608
A:Accession: AE0636
A>Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-495 <PAR>
A:Cross-references: UNIPARC:UPI0000059FC8; GB:AL513382; PIDN:CAD08272.1; PID:g16502319;
C:Genetics:
A:Gene: STY1185
C:Superfamily: cardiolipin synthase

Alignment Scores:
Pred. NO.: 1.92e-56 Length: 495
Score: 1089.50 Matches: 221
Percent Similarity: 63.9% Conservative: 77
Best Local Similarity: 47.4% Mismatches: 147
Query Match: 38.2% Indels: 21
DB: 2 Gaps: 5

US-10-665-990A-13 (1-1561) x AE0636 (1-495)

Qy 200 GGGCTATCCGACATCTACTCTCTCGAGCAACCCCAAGAAGCCCTTGGCGCCCGCGCGCC 259
Db 39 GlyGluCysGlyIleLeuAlaLeuAsnSerLeuAspAlaPheAlaAraGlyTyrArg 58
Qy 260 CTATCGAATCTCGCGAACACAGCCTCGATTGTGAATACTACATTTGGCGCAACGACATT 319
Db 59 LeuThrGluMetAlaAlaAraGlyLeuAspValGlnTyrIleTrpGluAspAspMet 78
Qy 320 TCCGGCAGGCTGTGTTTCAACCTCATGTACTCTCCGCAAGACGCGCGCTGGCGGTACGC 379

430 AlaThrLeuIleHisLysArgPheThrGlnSerGlnArgAspAlaAlaTrpGlnLeuArg 449

1430 CTCGACAAACACAAACCCGCTGCAATGGCAGCAT-----CCCGCCACCCGAAACACCTAC 1483

450 LeuAspArgTrpGlyArgIleAsnTrpIleAspArgGlnGluGluGluValLeu 469

1484 CCGAACCAACCCGAAAGCCAAACTTTGGAAACGATCGCGCAAAATCCTATCCCTGCTG 1543

470 LysLysGluProAlaThrArgPheTrpGlnArgValLeuValArgLeuAlaAlaIleLeu 489

1544 CCCATCGAAGGTTTATTA 1561

490 ProValGluTrpLeuLeu 495

RESULT 7

G97556

hypothetical 55.9K protein in csge-mdog intergenic region [imported] - Agrobacterium tum

C:Species: Agrobacterium tumefaciens

C:Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 09-Jul-2004

C:Accession: G97556

R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldman, A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.; Science 294, 2323-2328, 2001

A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum

A:Reference number: A97359; MUID:21608551; PMID:11743194

A:Accession: G97556

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-518 <KUR>

A:Cross-references: UNIPROT:Q8UEX3; UNIPARC:UPI000000D1C0B; GB:AE007869; PIDN:AAK87408.1

C:Genetics:

A:Gene: AGR_C_3007

A:Map position: circular chromosome

C:Superfamily: cardiolipin synthase

Alignment Scores:

Pred. No.: 7.12e-40 Length: 518

Score: 807.00 Matches: 204

Percent Similarity: 54.6% Conservative: 85

Best Local Similarity: 38.6% Mismatches: 196

Query Match: 28.3% Indels: 44

DB: 2 Gaps: 16

US-10-865-990A-13 (1-1561) x G97556 (1-518)

QY 53 CTCATTTCCTTTTATGCCCTCTCTCTGTTTCATGTTCTTTCATGTTGCCCCACCTGGAA 112

DB 8 IleIleIleLeuLeuMetLeuIleGlyProSerLeuPheValIleGlyLysGlnArg 27

QY 113 GAACGGACGGAAGCCGTCATTTTCAATCTTCCAAACCTCTCTCTCGGACCAACATCCTG 172

DB 28 GluLysAlaIleProLysArgProSerThrAlaLeuProValThrGluAspGluThrAla 47

QY 173 CAATCCGGACACCCCTCAT-AAAACGGGCTATCCGAC-----ATCTACCTG 220

DB 48 LeuAspArgHisTrpGlnSerIleArgAsnGlyTrpAsnGluLysAsnAlaLeuCysLeu 67

QY 221 CTGACGACCCCCACGAAGCCCTTGGCCCGCGCCCTTATCGAATCTGCGGACAC 280

DB 68 LeuHisSerAsnLeuAspAlaPheAlaValArgValAlaAlaAlaArgAlaAlaGlyArg 87

QY 281 AGCCTCGATTGCAATACTACATTTGCGCAACGACATTTCCGGCAGGCTGCTGTTCAAC 340

DB 88 SerLeuAspLeuMetTyrTyrMetTrpAsnAspLeuThrGlyArgLeuMetMetArg 107

QY 341 CTCATGTACCTTCCGCGAGAACCGCGCGTGCAGCTGCTGTGTGGACGACAAAC 400

DB 108 GluValIleAlaAlaAspArgGlyValArgValArgLeuLeuAspLeuGly 127

QY 401 ACGCGCGGTGGACGATCTCTGCTCGCCTCGACGACCATCCCAATATCGAAGTCCG 460

DB 128 ValSerMetSerAspArgIlePheHisAlaIleAspSerHisProAsnIleGluLeuArg 147

```

Db      490 LysileTyrArgArgGluProAspAlaAlaileSerArgilleAlaGlyileMet 509
QY      1535 TCCTGCTGCCCATCAAGGTTTATTA 1561
Db      510 ArgValLeuProLeuGluSerGlnLeu 518

RESULT 8
AB2777
phospholipase D family protein [imported] - Agrobacterium tumefaciens (strain C58, Dupont)
C:Species: Agrobacterium tumefaciens
C:Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 09-Jul-2004
C:Accession: AB2777
R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, I.;
erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutyavin, T.; Levy, R.; Li, M.; McClell
; Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,
ster, E.W.
A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A:Reference number: AB2577; PMID:21608550; PMID:11743193
A:Accession: AB2777
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-518 <KUR>
A:Cross-references: UNIPROT:Q8UEX3; UNIPARC:UPI00000D1C0B; GB:AE008688; PIDN:AAL42632.1;
A:Experimental source: strain C58 (Dupont)
C:Genetics:
A:Gene: Atul630
A:Map position: circular chromosome
C:Superfamily: cardiolipin synthase

Alignment Scores:
Pred. No.: 7,12e-40 Length: 518
Score: 807.00 Matches: 204
Percent Similarity: 54.6% Conservative: 85
Best Local Similarity: 38.6% Mismatches: 196
Query Match: 28.3% Indels: 44
DB: 2 Gaps: 16

US-10-665-990A-13 (1-1561) x AB2777 (1-518)

QY      53 CTCATTCCCTTTATGCTCTCTCTGTTTCATGTTCTTCATGTTGCCCCCATCGGAA 112
Db      8 IleIleIleLeuLeuMetLeuIleGlyProSerLeuPheValIleGlyLysGlnArg 27
QY      113 GAACGGACGGAAAGCGTCATTTCATTAATCTTCAAAACCTGTCCTCTCGACAAACATCCTG 172
Db      28 GluLysAlaIleProLysArgProSerThrAlaLeuProValThrGluAspGluThrAla 47
QY      173 CAAATCCGGCACACCCCTCAT---AACACGGGCTATCCGAC-----ATCTACCTG 220
Db      48 LeuAspArgHisTrpGlnSerIleArgAsnGlyTrpAsnGluLysAsnAlaLeuCysLeu 67
QY      221 CTCGACGACCCCAAGACCTTCGCGCGCGCGCGCCCTTATCGAATCTCGCCGAAC 280
Db      68 LeuHisSerAsnLeuAspAlaPheAlaValArgValAlaAlaAlaArgAlaGlyArg 87
QY      281 AGCCTCGATTTCGAATACATACATTTGGCGCAACGACATTTCCGGCAGGCTGCTGTTCAAC 340
Db      88 SerLeuAspLeuMetTyrTrpMetTrpAsnAlaAspLeuThrGlyArgLeuMetMetArg 107
QY      341 CTCATGTACTTTCGCGCAGAACGCGCGTGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 400
Db      108 GluValIleAlaAlaAlaAspArgGlyValArgValArgLeuLeuLeuAspAspLeuGly 127
QY      401 ACGCGGGGTTGGAGATCTCTGCTCGCCTCGACGACCATCCCAATATCAAGTGGC 460
Db      128 ValSerMetSerArgPheHisAlaIleAspSerHisProAsnIleGluLeuArg 147
QY      461 CTGTTCAACCCCTTCGTCCTAGCAAAA-----TGGCGGGCACTCGGCTACCTGACC 511
Db      148 LeuPheAsnProThrArgAlaArgGluAsnIleLeuHisArgSerLeuValLeu 167
```

```

QY      512 GACTTCCCTCCGCTCAACCCGCGCATGCACAACAATCTTTTACCGCGCAACAACCGCGCC 571
Db      168 ArgPheArgSerValAsnArgArgMetHisAsnLysAlaTrpIleAlaAspGlyArgAla 187
QY      572 ACCATCTCGGGGAGCGCAATATCGGACGAATACTTCAAGTCCGATCGGAGACACCGTT 631
Db      188 ValIleValGlyGlyArgAsnIleGlyAspAlaTyrPheAspAlaAlaGluArgAlaAsn 207
QY      632 TTCGCGGACCTTGGACATCTCTCCGCGCGAGCGTCGTCGCGAGATGATTCACGACATTC 691
Db      208 PheHisAspPheAspIleLeuGlyPheGlyGlyIleValAlaAspAlaThrGluIlePhe 227
QY      692 GACCGCTACTCGGCAAGCATTCGCCCAAC-----GCCACG 730
Db      228 AspAspTyrTrp-----AsnSerAlaValSerValProValArgSerLeuLeuAlaArg 245
QY      731 CGCATCATCCGCGCGGCAACATCGGCAAGGTTCTTCAAGCACTCGGATACAACGACGAA 790
Db      246 ArgProAsnLysLeuAlaLysLeuArgArgGluLeuAspAlaLeuProGlnSerGlu--- 264
QY      791 ACATCCAGACACGCGCTCTCTCGCTTACCGCGAAACCGTCGAACAGTCCGCCCTCTACCAA 850
Db      265 -----AlaAlaArgProTyrLeuGluArgValGluSerGln-----TyrGly 278
QY      851 AAA-----ATACAGACGGAGCGATCGACTGGCAGAGCGTCCAAACCGCG---CTG 898
Db      279 ArgAspHisPheLeuMetSerAspArgLeuHisTrp-----ValAspThrAlaAspVal 296
QY      899 ATCAGCAGACACCCCTGCAGAAAGGACTCGACCGCGCAACCGCAACCGCG---ATT 952
Db      297 LeuAlaAspProProGluLysAlaAlaGlyLys---ArgArgLysGlyHisAsnPheLeu 315
QY      953 GCGCGGAGGCTGCAAGACGCGCTCAACACAGCCCGCAAAAGCGTCTATCTGGTTTCACCC 1012
Db      316 MetGluSerLeuLeuProLeuMetGlnAlaAlaGlyGluSerLeuHisIleThrSerPro 335
QY      1013 TATTTCGTCCTACAAAATCCGGCACAGACGACCTGGCAAACTGGTCGAGGACGGCATA 1072
Db      336 TyrPheIleProGlyLysGlnGlyValGluIlePheLeuAspLeuAlaGluArgGlyVal 355
QY      1073 GACGTTACGCTTCGACCACTCGCTACAGCGCGACGCTTGGCGCGTCCATTCGCGC 1132
Db      356 SerLeuAlaIleLeuThrAsnSerLeuAlaAlaThrAspValAlaAlaValHisAlaGly 375
QY      1133 TACGTCAATACCGAAACCGCTGCTCAAGCCGCGCATCAAACTCTACAGCTGCAACCC 1192
Db      376 TyrAlaArgTyrArgLysProLeuLeuSerGlyGlyValArgLeuHisGluLeuArg--- 394
QY      1193 AACCATGCGTCCCGCCCAAAAGACAAAGGC-----CTGACCGGAGCTCCGTA 1243
Db      395 -----SerGlnAlaAspGlnGlySerPheThrLeuArgGlySerGlyGln 409
QY      1244 ACCGCTGCTATGCCAAAACCTTCATTGTGACGCGCAACGATCTTCTCGGCTCATTC 1303
Db      410 AlaSerLeuHisThrLysAlaPheThrArgAspGlyGluThrGlyTyrIleGlySerLeu 429
QY      1304 AACCTCGACCCCGTTCGCGACGCGCTCAATACCGAAATGGCGCTGCTCATCGAAAGCCCC 1363
Db      430 AsnPheAspProArgSerAlaSerLeuAsnThrGluMetGlyValValPheAsnSerAla 449
QY      1364 AAAATCGCAGAACAGATGGAGGCGACCTCGCCGATACCACACCGAA---TACGCCATC 1420
Db      450 ProLeuValAlaArgMetAspGluIlePheAlaGluIleArgArgThrMetSerPhe 469
QY      1421 CGCGTTACCTCGCAAAACACACCGCTCGCAATGG-----CACGATCCCGCCACCGA 1474
Db      470 GluLeuAspIleAspSerAlaAsnArgIleValTrpMetThrGluGluArgGlyGlnPro 489
QY      1475 AAAACCTTACCGAACGAAACCGGACCGCAAACTTTGGAAACGATCGCGCGCAAAATCTTA 1534
Db      490 LysIleTyrArgArgGluProAspAlaAlaIleSerArgGlyIleAlaGlyIleMet 509
```

177	HisAsnLysLeuTrpLeuAlaAspGlyThrAlaAlaIleValGlyArgAsnLeuGly	196
599	GACCAATACTTCAAAGTCGGTGAGACACCGTTTTTCGCCGACCTGCAGATCTCTCGCCACC	658
197	AspGluTyrPheAsnAlaLysProGluMetAsnPheThrAspLeuAspLeuGlyVal	216
659	GGCAGCGTCGTCGGCGAAGTATCGCACACATTCGACCGCTACTGGGCAAGCCATTCGCGC	718
217	GlyProIleAlaAsnGlnLeuSerHisSerPheAspGlnTyrTrpAsnSerAlaIleSer	236
719	CACAACGCCACG-----CGCATCATCTCCGACGCGCGCAACATCGCGAAGGGTCTT	766
237	ArgProIleGluAspPheLeuTrpArgAlaProTyrProGlyGluLeu-----	252
767	CAACGACTCGGATACACGACGAAACATCCAGACACGCGCTCTCGGTAC-----CGC	820
253	-----GluSerAlaArgLysLeuGlnArgTyrLeuArgLys	265
821	GAACACCGTCGAACAGTCG-----CCCTCTACCAAAAAAATACAGACGGGACGC	868
266	GluSerValLysGluSerGlyTyrIleArgHisLeuPheAspArgGlyAspGlnProArg	285
869	ATC---GACTGGCAGAGC-----GTCCAAACCCGCGCTGATCAGCAGACCCCT	913
286	LeuGlyAsnTrpLeuGluAsnLeuThrTrpAlaArgAlaGluAlaIleTrpAspAlaPro	305
914	GCAAAAGGACTCGACCGCGCCGCCGCAAAACGCGCG-----ATTGCCGGAGGCTGC	967
306	LeuLysValLeuSerArgGlyGluProAspProHisLeuLeuLeuSerProHisLeuAla	325
968	GACGCGCTCAACACGCGCGAAAAAGCGTCTATCTGGTTTCACCCCTATTTCGTCCTACA	102
326	GlyLeuPheLysGlyValGlnLysGluLeuLeuValSerAlaTyrPheValProAla	345
1028	AAATCCGGCACAGCAGCGCTGGCAAAATCGTGTGCAGGACGGCATAGAGTTTACCGTCCTG	108
346	LysAspGlyLeuAsnTyrLeuThrGlyLysAlaAspSerGlyValArgValArgLeuLeu	365
1088	ACCAACTCGGTACAGCGCGCAGCGTTCGCCGCGCTCCATTCCGGGTACGTCAAAATACCGA	114
366	ThrAsnSerLeuGluAlaThrAspValProAlaValHisAlaGlyTyrAlaProTyrArg	385
1148	AAACCGCTGTCAAAGCGCGCATCAAACCTCTACGAGCTGCAACCAACCATCGCGTCCCC	120
386	MetAlaLeuLeuGluHisGlyValLysLeuTyrGluLeuArgAlaAsnProAspGlnPro	405
1208	GCCACA---AAAGACAAAGCGCTGACCGGCAGCTCCGTAAACCGCCTGCATGCCAAAACC	126
406	LeuSerGlyAlaProTyrArgLeuHisGlySerSerSerAlaSerLeuHisSerLysAla	425
1265	TTCAATTGTGACGCGCAACGCATCTTATTCGGCTCATTTCAACCTCGACCCCGTTCGCGCA	132
426	MetValPheAspArgArgLysValPheIleGlySerPheAsnPheAspProArgSerIle	445
1325	CGGCTCAATACCCGAAATGGCGTTCGTATCAAAAGCCCAAAATCGCAGAACAGATGGAG	138
446	LeuTrpAsnThrGluValGlyValIleValAspSerProLeuLeuAlaGluGlnValArg	465
1385	CGC---ACCCCTGCGCGATACACACCCGGAATACGCCCTTACCGGTTTACCTCGACAAACAC	144
466	GlnLeuAlaLeuGluGlyMetAlaProSerValSerTyrGlnValArgIleAspArgSer	485
1442	AAC-----CGCCTGCAATGCCAGAT-----CCCGCCACCCCGAAAAACCTTACCGC	148
486	GlySerArgProLysLeuValTrpIleAspGluArgAspGlyArgAlaGlnValLeuArg	505
1487	AACGAACCCGAAGCCAAACTTTGGAAAGCATCGCCGCAAAAATCCTATCCTCGTGTGCC	154
506	HisGluPro---GlySerLeuTrpArgArgLeuAsnAlaTrpValAlaGlyMetIleGly	524
1547	ATCGAAGGTTTATTA	1561

Qy	914	GC AAAAGGAGCTGCAGCCGCGCCGCGCAAAACCGCGC-----ATTGCGGGGAGGCTGC AA	967
Db	306	LeuLysValLeuSerArgGlyGluProAspProHisLeuLeuLeuSerProHisLeuAla	325
Qy	968	GACCGCTCAACAGCCCGGAAAAAGGCTCTATCTGGTTTCACCCCTATTTCCTCCCTACA	1021
Db	326	GlyLeuPhelysGlyValGlnLysGluLeuLeuValSerAlaTyrPheValProAla	345
Qy	1028	AAATCCGGCACAGACGCACTGGCAAAACTGTGTGCAGCAGCGCATAGAGTTTACCGTCCTG	1088
Db	346	LysAspGlyLeuAsnTyrLeuThrGlyLysAlaAspSerGlyValArgValArgLeuLeu	365
Qy	1088	ACCAACTCGCTACAGCGCAGCGAGTTCCCGCGCTCCATTCCGGGTACGTCAAATACCGA	114
Db	366	ThrAsnSerLeuGluAlaThrAspValProAlaValHisAlaGlyTyrAlaProTyrArg	385
Qy	1148	AAACCGCTGTCTAAAGCGCGCATCAAACTCTACGAGCTGCACCAACCCACCATGCCGTCGCC	120
Db	386	MetAlaLeuLeuGluHisGlyValLysLeuTyrGluLeuArgAlaAsnProAspGlnPro	405
Qy	1208	GCACA---AAAGACAAAGCGCTACCGGAGCTCCGTAAACGAGCTGCATGCCAAAACC	126
Db	406	LeuSerGlyAlaProTyrArgLeuHisGlySerSerSerAlaSerLeuHisSerLysAla	425
Qy	1265	TTCAATCTGCAGCGCAACGCACTTCATTCGGCTCATTCACACTCGACCCCGTTCGCGCA	132
Db	426	MetValPheAspArgArgLysValPheIleGlySerPheAsnPheAspProArgSerIle	445
Qy	1325	CGGCTCAATACCGAAATGGCGGTGCTATCGAAGCCCAAAATCGCGAAGACAGATGGAG	138
Db	446	LeuTrpAsnThrGluValGlyValIleValAspSerProLeuLeuAlaGluGlnValArg	465
Qy	1385	CGC---ACCCTCGCGCATACACACCCGGAATACGCCTACCGCTTACCTCGCAAAACAC	144
Db	466	GlnLeuAlaLeuGluGlyMetAlaProSerValSerTyrGlnValArgIleAspArgSer	485
Qy	1442	AAC-----CGCTGCAATGCAAGAT-----CCCGCACCCCGGAAAAACCTACCGG	148
Db	486	GlySerArgProLysLeuValTrpIleAspGluArgAspGlyArgAlaGlnValLeuLeuArg	505
Qy	1487	AACGAACCCGAAGCCAAACTTGAAGACGATCCCGCAAAAATCTTATCTCCGTGTCGCC	154
Db	506	HisGluPro---GlySerLeuTrpArgArgLeuAsnAlaTrpValAlaGlyMetIleGly	524
Qy	1547	ATCGAAGGTTTATTA	1561

Db 525 LeuGluLysMetLeu 529

RESULT 10

probable cardiolipin synthase - Helicobacter pylori (strain J99)

C.Species: Helicobacter pylori

A.Variety: strain J99

C.Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 09-Jul-2004

C.Accession: E71963

R.Alm, R.A.; Ling, L.S.L.; Moir, D.T.; King, B.L.; Brown, E.D.; Doig, P.C.; Smith, D.R.; Ives, C.; Gibson, R.; Merberg, D.; Mills, S.D.; Jiang, Q.; Taylor, D.E.; Vovis, G.F.; Nature 397, 176-180, 1999

A.Title: Genomic sequence comparison of two unrelated isolates of the human gastric path

A.Reference number: A71800; MUID:99120557; PMID:9923682

A.Accession: E71963

A.Status: preliminary

A.Molecule type: DNA

A.Residues: 1-502 <ARN>

A.Cross-references: UNIPROT:Q9ZMP2; UNIPARC:UPI000013942E; GB:AE001456; GB:AE001439; NID

A.Experimental source: strain J99

C.Genetics:

C.Superfamily: cardiolipin synthase

Alignment Scores:

Pred. No.:	1.3e-29	Length:	502
Score:	632.00	Matches:	161
Percent Similarity:	51.7%	Conservative:	93
Best Local Similarity:	32.8%	Mismatches:	181
Query Match:	22.2%	Indels:	56
Db:	2	Gaps:	17

US-10-665-990A-13 (1-1561) x E71963 (1-502)

QY 176 ATCCGGCACACCCCTCATAAACAGGGCTATCCGACATCTAC----- 217

Db 30 lIeSerTyrAspProTyrThrThrThrleGlySerLeuTyrAlaLysAsnLeuLysGlu 49

QY 218 -----CTGCTCGACGACCCCGACAGCCCTTGC----- 253

Db 50 AsnProLysHisSerAlaAlaileLeuLeuGluAspGlyPheAspAlaLeuHisArg 69

QY 254 GCCGCCCTATCAATCTGCCAAACACAGCTCGATTGTCATACATACATTTGGCGCAAC 313

Db 70 ValGlyLeuileArgMetSerGlnLysSerIleAspMetGlnThrTyrileTyrLysAsn 89

QY 314 GACATTTCCGGCAGCGTCTGTTCACCTCATGTACCTTGCCTGCGCAGAACCGCGCTGCGC 373

Db 90 AspLeuSerSerGlnValileAlaLysGluLeuLeuAsnAlaAlaAsnArgLysVallys 109

QY 374 GTACGCTGCTCTTGGACGACAAACACAGCGCGGGTTGGAC-----GATCTC 421

Db 110 ValArgileLeuLeuAspAsn-----GlyLeuAspSerAspPheSerAspile 126

QY 422 CTGCTCGCCTCGACAGCATCCCAATATCGAAGTGGCGCTTCAACCCCTTCCTCCTA 481

Db 127 MetLeu---LeuAsnPheHisLysAsnileGluValLysilePheAsnProTyrTyrile 145

QY 482 CGCAAAATGGCGCAGCTCGGCTAC-----CTGACCGACTTCCCGCCTCAACCGC 532

Db 146 ArgAsn---LysGlyLeuArgTyrPheGluMetLeuAlaAspTyrGluArgileLysLys 164

QY 533 CGCATGCAACAAATCTTTTACCGCGCAACACCGCGCACCATCTCGGCGGAGCGCAAT 592

Db 165 ArgMetHisAsnLysLeuPheileValAspAsnPheAlaValileileGlyGlyArgAsn 184

QY 593 ATCCGGCGCATATCTTCAAGTCGGTGAGGACCGTTTTCGCGCAGCTGACATCTC 652

Db 185 lIeGlyAspAsnTyrPheAspAsnAspLeuAspThrAsnPheLeuAspLeuAspAlaLeu 204

QY 653 GCCACCGGCGAGCGTCTGTCGGCGAAGTATCGCACGACTTTCGACCGCTACTGGCGAAGCCAT 712

Db 205 PhePheGlyGlyValAlaSerLysAlaLysGluSerPheGluAsnTyrTrpArgPheHis 224

QY 713 TCCGCCCAACAGCCGACGCGCATCATCCGAGC-----GGC 748

Db 225 ArgSerileProValSer---LeuLeuArgThrHisLysArgLeuLysAsnAsnVallys 243

QY 749 AACATCGCAAGGCTCTTCAGCACTCGATACACGACGAAACATTCAGACACGCGCTC 808

Db 244 GluileAlaLysLeuHisGluLysileProleserAlaGluAspAlaAsnGluPheGlu 263

QY 809 CTGGCTACCCGCAAAACCGTCAACGCTCGCCCTCTACCAA---AAATACAGCGGA 865

Db 264 LysLysValAsnAspPheileGluArgPheGlnLysTyrGlnTyrProileTyrTyrGly 283

QY 866 CGCATCGACTGGCAGAGCGTCAAAACCGCTGTATCAGCGACACCCCTGCAAAAGGACTC 925

Db 284 -----AsnAlailePheLeuAlaAspLeuProAlaLys---ile 295

QY 926 GACCGGACCGCGCAACCGCGATTGCGGGAGGCTGCAAGAGCGCTCAACAGCGCC 985

Db 296 AspThrProLeuTyrSer---ProleLysileAlaPheGluLysAlaLeuLysAsnAla 314

QY 986 GAAAAAGCGTCTATCTGTTTTCACCTTATTTCTGTCCTACAAAATCCGCGCACAGCA 1045

Db 315 LysAspSerValPheileAlaSerSerTyrPheileProGlyLysLysileMetLysile 334

QY 1046 CTGGCAAACTGTGTGAGGACGGCATAGACTTACCTGCTCTGACCAACTCGTACAGGG 1105

Db 335 PheLysAsnGlnileSerLysGlyileGluLeuAsnileLeuThrAsnSerLysSer 354

QY 1106 ACCGAGCTTGGCGCGCTCATTTCCGGTACGTCAATACCGAAAACCCCTGCTCAAAGCC 1165

Db 355 ThrAspAlaileValValTyrGlyAlaTrpGluArgTyrArgAsnLysLeuValArgMet 374

QY 1166 GGCATCAAACTCTACGAGCTGCAACCCCAACCATGCCGTCCCGCCACAAAAGCAAGGC 1225

Db 375 GlyAlaAsnValTyrGluileArgAsnAspPhePheAsnArgGlnileLysGlyArg--- 393

QY 1226 CTGACCGGCGCTCGTNAACCGACTGATGCCAAAACCTTCATTGTGCGCGCAACGC 1285

Db 394 -----PheSerThrLysHisSerLeuHisGlyLysThrileValPheAspAlaLeu 411

QY 1286 ATCTTCATCGCTCATTCACACTCGACCCCGTTCGCGACCGCTCAATACCGAAATGGC 1345

Db 412 ThrLeuLeuGlySerPheAsnileAspProArgSerAlaTyrIleAsnThrGluSerAla 431

QY 1346 GTGCTCATCGAAAGCCCAAAATCGCAACAGATGGAGCGACCTCGCGCATACCA 1405

Db 432 ValLeuPheAspAsnProSerPheAlaLysArgValArgLeuSerLysLysAspHisAla 451

QY 1406 CCCGNATACGCTACCGGTTACCTCGACAAACACACCGCTGCAATGGCAGCATCCC 1465

Db 452 ---GlnGlnSerTrpHisLeuValLeuTyrArgHis---ArgValileTrpGlu----- 467

QY 1466 GCCACCCGAAAAACCTAC-----CCGAACGAACCCGAAACCAAACTTTGGAAA 1513

Db 468 AlaThrGluGluGlyileLeuileHisGluLysAsnSerProAspThrSerPhePheLeu 487

QY 1514 CGCATCGCGCAAAATCTTATCCCTGCTGCC 1546

Db 488 ArgLeuileLysGluTrpSerLysValLeuPro 498

RESULT 11

F64543

conserved hypothetical secreted protein HP0190 - Helicobacter pylori (strain 26695)

C.Species: Helicobacter pylori

C.Date: 09-Aug-1997 #sequence_revision 09-Aug-1997 #text_change 09-Jul-2004

C.Accession: F64543

R.Tomb, J.F.; White, O.; Kerlavage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R.D.; Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khaliak, H.G.; Glodek, A.; McKenrison, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.; Bowman, C.; Watthey, L. Nature 388, 539-547, 1997

A.Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karpk, P.D.; Smith, H.O.; Fraser, C. A>Title: The complete genome sequence of the gastric pathogen Helicobacter pylori.

Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasak
A;Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveira
M.; Tshako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z
A;Reference number: A59328
A;Contents: annotation
C;Genetics:
A;Gene: XFI087
C;Superfamily: cardiolipin synthase

Alignment Scores:

Pred. No.: 2,76e-18 Length: 652
Score: 439.00 Matches: 161
Percent Similarity: 40.1% Conservative: 69
Best local Similarity: 28.0% Mismatches: 200
Query Match: 15.4% Indels: 144
DB: 2 Gaps: 21

US-10-665-990A-13 (1-1561) x B82724 (1-652)

```
QY 125 AGCGTCAATTCATCTTCAATCTTCAAACTGTCTCTCTGGACAACATCTCTGCAATCCGGCAC 184
||||| : : ||| |||||
Db 61 SerArgAlaLeuSerGluSerThrProVal----- 70

QY 185 ACCCTCATACAAACGGCTCTCCGACATCTACCTGTCTGAGACGCCGCCAGAGCCCTT 244
||||| : : ||| |||||
Db 71 ThrPro-----LeuHisTyrValThrIleLeuAspLysGlyGluSerAlaLeu 86

QY 245 GCGCGCCGCGCGCCCTTATCGAATCTGCGGAACACAGCCTCGAATTCGAATCTACATTT 304
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 87 ValAlaArgIleAsnLeuIleArgSerAlaThrArgSerIleAspMetGlnThrTyrIle 106

QY 305 TGGCGCAACGACATTTCCGGCAGGTGCTGTTCAACTCATGTATCTTCCGCGAGAACGC 364
||||| ||||| ||||| : : : : |||||
Db 107 PheAspLysAspSerAlaArgLeuIleMetAspGluLeuLeuThrAlaAlaArg 126

QY 365 GCGCTGCGGTACGCTGCTGTGACGACACACACACGCGCGGTGGAGCATCTCTCG 424
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 127 GlyValArgValArgLeuIleAspGlnLeuSerAlaIleSerAspLeuAsnIleLeu 146

QY 425 CTCGCGCTC---GACAGCATCCCAATATCGAAGTGGCGCTGTTCAACCCCTCTCTCTA 481
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 147 GlyAlaLeuAlaGlyAlaHisValAsnPheGlnLeuArgIleTyrAsnProIlePheGly 166

QY 482 CGCAATGGCGGCGCTCGGCTAC-----CTGACCGACTTCCCC 520
: : : ||||| ||||| |||||
Db 167 Lys-----AlaLysLeuAsnTyrGlyAspTyrValAlaSerValLeuCysCysPheArg 184

QY 521 CGCCTCAACCGCGCATGCACACAAATCTTTTACCGCGCAACCGCGCCACCATATCTC 580
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 185 ArgPheAsnGlnArgMetHisAsnLysLeuLeuValIleAspGluMetIleGlyVal 204

QY 581 GCGGACGCAATATCGGCGCAAAATCTTCAAAGTCGGTGGAGACACCGTTTTCGCGCAC 640
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 205 GlyGlyArgAsnTyrGlnAspAspTyrTyrAspTyrAspLeuGluTyrAsnPheArgAsp 224

QY 641 CTGGACATCTCGCCACCGGAGCGTGTGCGGGAAGTATCCGACGACTTGGACCGCTAC 700
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 225 ArgAspValLeuValAlaGlyProAlaValLeuGlnMetAlaValAsnPheAspAlaPhe 244

QY 701 TGGGCAAGCCATCTCGCCACACGCCACGCGCATCTCCGCGGCAACATCGGCAAG 760
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 245 TrpAla-----AlaGluArgSerValProValGluCysLeuArgAspValGlyArg 261

QY 761 -----GGTCTTCAAGCACTCGGATACAAACGACGAAACATCCAGACACGCG 805
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 262 MetValLeuGlnAspGlyValProMetLeu-----ProSerAlaValPheHisPro 278

QY 806 CTCCTGCGC---TACGGGAACCGTCGAA---CAGTCGCCCTCTACCAAAA----- 853
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 279 LeuGluArgValGlnArgValThrAlaGluAlaAsnAspProAspTyrValLysArgThr 298

QY 854 -----ATACAGAGCGGACGCGCATCTGCGAGCGTCCAAACCCGCGCTG 898
: : : ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

```
Db 299 PheValAspProAlaLeuAlaValAsnLysValGlnTyr----- 311
QY 899 ATCAGCGACCCCTGCAAAAGAGACTCGACCGCGACCGCGCAAAACCG---CCGATTGCC 955
: : : ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 312 ValAlaAspLeuProGlnLysHisArgHisGluHisAsnAlaAsnProValSerIleGly 331
QY 956 GGGAGGCTGCAAGACGGCTCAAAACAGCCCCGAAAGAGCGTCTATCTGGTTTACCCTAT 1015
||||| ||||| : : : ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 332 ProArgLeuAspSerLeuIleSerAsnAlaArgHisGluValIleLeuGlnThrProTyr 351
QY 1016 TTCGTCTCCATAAAATCCGGCACAGCAGCACTGCGAAACCTG-----GTGCGAGCGC 1069
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 352 LeuValLeuSerLysProAlaLeuAsnIlePheArgLeuAsnArgThrGlnAspLys 371
QY 1070 ATAGACGTTTACCGTCTGACCACTCGCTACAGCGACCGACGCTTGGCGCGCTCCATTCC 1129
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 372 ProArgValValValThrAsnSerLeuAlaIleThrAspAsnProIleValTyrAla 391
QY 1130 GGCTACGTCAATAC---CGAAACCGCTGCTCAAGCGCGCATCAAACTCTACGAGCTG 1186
||||| ||||| ||||| : : : ||||| ||||| ||||| ||||| ||||| |||||
Db 392 LeuPheTyrLysTyrLysArgAsnMetArgAspLeuGlyPheAspIleTyrGluTyr 411
QY 1187 CAACCC----- 1192
Db 412 LysProPheProGlnAsnProProIleAspLeuThrGlyValValProMetAspGlyTrp 431
QY 1192 ----- 1192
Db 432 AsnAsnAspSerLeuGlnThrArgGlnAspLeuAlaAlaLysValSerAspAsp 451
QY 1193 -----AACCATGCGTCCCGCCACAAAGACAAAGCGCTGACCGGAGCTCCGTA 1243
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 452 SerAspAsnAlaArgGlnLeuProSerLeuGlnAsnLysGlyGlnValAspArgVal 471
QY 1243 ----- 1243
Db 472 LeuArgThrGluThrArgProPheLeuGlyIleHisThrValAsnLysProLeuPro 491
QY 1244 -----ACCAGCTGATGCGCAAAACCTTCATTGTGTGACGGCAAA 1282
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 492 ValThrArgProGlyAlaArgMetGlyLeuHisAlaLysSerLeuValValAspArgLys 511
QY 1283 CGCATCTTCATCGGCTCATTAACCTCGACCCCGTTCGCGACGCTCAATCCGAAATG 1342
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 512 ValGlyValIleGlyThrHisAsnPheAspProArgGlyGluSerTyrAsnThrGluAla 531
QY 1343 GCGCTGCTCATCGAAAGCCCAAAATCGCAAAACAGATGGAGCGCACCCCTC---GCCGAT 1399
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 532 ValValValIleGluAspProAlaPheAlaArgLeuLeuAlaSerSerIleGluGlyAsp 551
QY 1400 ACCACACCGCAATACGCTACCGGCTTACCCTCGAACAACACACCGCTGCAATGGCAC 1459
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 552 MetAspProGlyAsnAlaTrpValVal----- 560
QY 1460 GATCCCGCACCCGAAACCTACCG-----AACGAACCCGAAAGCAAA 1504
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 561 -----AlaProArgLysLysLeuProGlyLeuTyrLysLeuAsnTyrSerValGlyLys 578
QY 1505 CTTTGGAAACGATCGCGCAAAATCCTCCTCTGCTGCC 1546
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 579 LeuSerGlu-----AlaLeuProValLeuAspLeuTrpPro 590
```

RESULT 13

A43932

mucin 2 precursor, intestinal - human (fragments)

N;Alternate names: mucin SMUC-41

C;Species: Homo sapiens (man)

C;Date: 10-Mar-1993 #sequence revision 12-Apr-1996 #text change 31-Dec-2004

C;Accession: A49963; A45106; B45106; A43932; B35332; A61257; P00328; P00329

J;Gum Jr., J.R.; Hicks, J.W.; Toribara, N.W.; Siddiki, B.; Kim, Y.S.

J;Biol. Chem. 269, 2440-2446, 1994

A;Title: Molecular cloning of human intestinal mucin (MUC2) cDNA. Identification of the

A;Reference number: A49963; MUID:94132002; PMID:8300571

A:Accession: A49963
A:Molecule type: mRNA
A:Residues: 1-639 <GUI>
A:Cross-references: UNIPROT:Q02817; UNIPROT:Q14884; UNIPARC:UPI0000177AE0; GB:L21998
R:Gum Jr., J.R.; Hicks, J.W.; Toribara, N.W.; Roche, E.M.; Lagace, R.E.; Kim, Y.S.
J. Biol. Chem. 267, 21375-21383, 1992
A:Title: The human MUC2 intestinal mucin has cysteine-rich subdomains located both upstr
A:Reference number: A45106; MUID:93016075; PMID:1400449
A:Accession: A45106
A:Status: not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 628-1895 <GUG>
A:Cross-references: UNIPARC:UPI000016AB12; GB:M94131; NID:9186395; PIDN:AAA59163.1; PID:
A:Note: sequence extracted from NCBI backbone (NCBIP:116706)
A:Accession: B45106
A:Status: not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 2037-3020 <GUG>
A:Cross-references: UNIPARC:UPI000016AB13; GB:M94132; NID:9186397; PIDN:AAA59164.1; PID:
A:Experimental source: colon
A:Note: sequence extracted from NCBI backbone (NCBIP:116698)
R:Toribara, N.W.; Gum Jr., J.R.; Culhane, P.J.; Lagace, R.E.; Hicks, J.W.; Petersen, G.M.
J. Clin. Invest. 88, 1005-1013, 1991
A:Title: MUC-2 human small intestinal mucin gene structure. Repeated arrays and polymor
A:Reference number: A43932; MUID:91358717; PMID:1885763
A:Accession: A43932
A:Molecule type: DNA
A:Residues: 1343-1350, 'L', 1352-1411, 'S', 1413-1448, 'P', 1450-1503, 'T', 1505-1915 <TOR>
A:Cross-references: UNIPARC:UPI000016AD85; GB:M74027; NID:9188863; PIDN:AAA59875.1; PID:
A:Note: sequence inconsistent with the nucleotide translation
A:Note: sequence extracted from NCBI backbone (NCBIP:55749)
R:Gum, J.R.; Byrd, J.C.; Hicks, J.W.; Toribara, N.W.; Lampert, D.T.A.; Kim, Y.S.
J. Biol. Chem. 264, 6480-6487, 1989
A:Title: Molecular cloning of human intestinal mucin cDNAs. Sequence analysis and eviden
A:Reference number: A33532; MUID:89197956; PMID:2703501
A:Accession: B33532
A:Molecule type: mRNA
A:Residues: 1916-2193 <GUA>
A:Cross-references: UNIPARC:UPI00000731C8; GB:M22405; NID:9188873; PIDN:AAA36334.1; PID:
A:Experimental source: intestine
R:Jany, B.H.; Gallup, M.W.; Yan, P.S.; Gum, J.R.; Kim, Y.S.; Basbaum, C.B.
J. Clin. Invest. 87, 77-82, 1991
A:Title: Human bronchus and intestine express the same mucin gene.
A:Reference number: A61257; MUID:91086481; PMID:1985113
A:Accession: A61257
A:Status: not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 'T', 1925-1948, 'TTS', 1952-1954 <JAN>
A:Cross-references: UNIPARC:UPI0000177AE1
R:Xu, G.; Huan, L.; Khatri, I.; Sajjan, U.S.; McCool, D.; Wang, D.; Jones, C.; Forstner,
Biochem. Biophys. Res. Commun. 183, 821-828, 1992
A:Title: Human intestinal mucin-like protein (MUP) is homologous with rat MLP in the C-t
A:Reference number: FQ0328; MUID:92198477; PMID:1550588
A:Accession: FQ0328
A:Molecule type: mRNA
A:Residues: 2328-2468 <XUG>
A:Cross-references: UNIPARC:UPI0000177AE2; GB:M86523
A:Experimental source: small intestine
A:Accession: FQ0329
A:Molecule type: protein
A:Residues: 2328-2342, 'K', 2344-2354 <XUG1>
A:Cross-references: UNIPARC:UPI0000177AE3
C:Genetics:
A:Gene: GDB:MUC2
A:Cross-references: GDB:120203; OMIM:158370
A:Map position: lip15.5-lip15.5
C:Superfamily: von Willebrand factor type C repeat homology
C:Keywords: glycoprotein; intestine; tandem repeat
F:2766-2834/Domain: von Willebrand factor type C repeat homology <VWC>
Alignment Scores:
Pred. No.: 3.17e-12 Length: 3020

Score: 336.50 Matches: 141
Percent Similarity: 33.6% Conservative: 55
Best Local Similarity: 24.1% Mismatches: 198
Query Match: 11.8% Indels: 190
DB: 2 Gaps: 22
US-10-665-990A-13 (1-1561) x A43932 (1-3020)
QY 119 ACGGAAAGCCGCTCATTTCAAT-----ACTTCCAAACCTGCTCTCTGGACAAAC 166
Ddb 1248 ThrValGluLysHisPheAsnIleCysSerIleThrThrArgPro-SerThrLeuThrTh 1267
QY 167 ATCTGTCAGAAATCCGGCAGCACCCCTCATAAACAACGGGTATCCGACATCTACTGCTGCAC 226
Ddb 1267 rPheThrThrIleThrLeuProThrThrProThrSerPheThrThrThrThrThrTh 1287
QY 227 GACCCCAACGAGCCCTTGGCGCGCGCCCTTATCGAATCTGCGCAACACGACGCTC 286
Ddb 1287 rThrProThr-----SerSerTh 1293
QY 287 GATTTGCAATACATTTGGCGCAACGACATTTTCGGCAGGCTGCTGTTCAACCTCA-- 344
Ddb 1293 rValLeuSerThrThr-----ProLysLeuCysCysLeuTrpSerAs 1307
QY 344 ----- 344
Ddb 1307 pTrpIleAsnGluAspHisProSerSerGlySerAspAspGlyAspArgGluProPheAs 1327
QY 345 -----TGTACTTGGCGGAGAACGGCGGTGGCGTACGCC----- 380
Ddb 1327 pGlyValCysGlyAlaProGluAspIleGluCysArgSerValLysAspProHisLeuSe 1347
QY 380 ----- 380
Ddb 1347 rLeuGluGlnHisGlyGlnLysValGlnCysAspValSerValGlyPheIleCysLysAs 1367
QY 380 ----- 380
Ddb 1367 nGluAspGlnPheGlyAsnGlyProPheGlyLeuCysTyAspTyryLysIleArgValas 1387
QY 381 -TGCTGTGG-----ACGACAACAAC 400
Ddb 1387 nCysCysTrpProMetAspLysCysIleThrThrProSerProProThrThrThrProSe 1407
QY 401 ACGCGCGGGTGGAGCATCTCTGCTCGCCTCGACAGCCATCCCATATATCAAGTGGCG 460
Ddb 1407 rProProProThrThrThrLeuProProThrThrPro----- 1422
QY 461 CTGTTCAACCCCTTCGTCCTACGCAAAATGGCGGCACTCGGCTACCTGACCGACTTCCCC 520
Ddb 1423 -----SerProProThrThrThrThrThrThrProProProThrThrThrProSerProPr 1441
QY 521 CGCTCAACCGCGCATGCAACAATCTTTACCGCGGACACACCGCGGACCATCTACTC 580
Ddb 1441 oIleThrThr-----ThrThrThrProLeuProThrThrThrProSerPro----- 1456
QY 581 GCGCGGCAATATCGCGAGCAATCTTCAAACTCGTGAGGACACCGTTTTCGCGCAC 640
Ddb 1457 -----ProIleSerThrThrThrProProProProThrThrThrThrThrThrPro 1474
QY 641 CTGACATCTCTGCCACCGCGAGCGTCTGCGGCGAGTATCGCACGACTTCGACG----- 695
Ddb 1474 rThrThrProSerProProThrThrThrThrThrThrThrThrThrThrThrThrThrPr 1494
QY 696 -----GCTACTGGGCAAGCCATTCGCCCAACAACGCCACGGCGCATCATCCGAGCGGC 748
Ddb 1494 oProProThrThrThrProSerProProMetThrThrProIleThrProProAlaSerTh 1514
QY 749 AACATCGGCAAGGCTCTTCAAGCACTCGGATACCAACGACGAAACATCCAGACGCGCTC 808
Ddb 1514 rThrThrLeu-----ProProThrThrThrProSerProProThrThrThrTh 1529
QY 809 CTGCGGTACCGGAAACCGTCGAAC-----AGTCGCGCCCTCTACCAAAAAATACAGACG 862


```
Db 404 uSerSerAlaProValThrSerSerThrThrThrThrThrThrThrThrThrThrThrThrThrThrSe 424
QY 536 ATGCACAAACAAAT-----CCTTTACCGCGGACAAACCGCGCCACATATCTCGC 583
Db 424 rSerThrThrGluSerSerAlaProValThrSerSerThrThrThrThrThrThrThrThrThrThr 444
QY 584 GGACGCAATATATCGGCGGACGAATACCTTCAAAGTCG-----GTGAGGACACCGTTTTCG-- 635
Db 444 aProValThrSerSerThrThrGluSerSerAlaProValProValProValProValProValSerSe 464
QY 635 ----- 635
Db 464 rThrThrGluSerSerAlaProValThrSerSerThrThrThrThrThrThrThrThrThrThrThr 484
QY 636 ---CGGACCTGGACATCTCGCCACCGCGCGTCGTCGCGGAAGTATCGCACACTTC 691
Db 484 oValProThrProSerSerThrThrThrGluSerSerAlaProValThrThrThrThrThrThrThr 504
QY 692 GACCGCTACTGGCGCAAGCATTCGCCCCACACA----- 722
Db 504 rThrGluSerSerAlaProValProValProValProValProValProValProValProValProVal 524
QY 723 -----ACGCCACCGGCATCATCCGACGCGCAACATCGCGCAAGGGTCTTCAA 769
Db 524 rAlaProAlaProThrProSerSerThrThrThrGluSerSerAlaProValThrSe 544
QY 770 GCAC-----TCGNATACACGACGAAACA 793
Db 544 rSerThrThrGluSerSerAlaProValProValProValProValProValProValProValProVal 564
QY 794 TCCAGACACGCGCTCTCGCTACCGCGAAACCGTCGAACAGTCGCCCTCTACCAAAA 853
Db 564 rSerSerThrProValThrSerSerThrThrThrGluSerSerAlaProValProValProValProVal 584
QY 854 ATACAGCGGACGATCGATCTGGCAGAGCGTCCAAACCG--CCTGATCAGCGACACC 910
Db 584 roSerSerSerThrThrGluSerSerAlaProValProValProValProValProValProValProVal 604
QY 911 CCTGAAAGGACTCGACCGCGCGCGCGCAACCGCGATTCGCGGGAGCTCGAAGC 970
Db 604 hrGluSerSerAlaProAlaProAlaProThrProSerSerSerThrThrThrGluSerSerAlaPr 624
QY 971 CGCGCTCAACACGCGCAAAAAGCGCTCTATCTGGTTTCACCCCTATTTCGTCCTCACAAA 1030
Db 624 laProValThrSerSerThrThrGluSerSerAlaProValProValProValProValProValProVal 643
QY 1031 TCCGGCACAGACGCACTGGCAAAACTGGTGAGGACGCGCATAGACGTTCACGTCCTGACC 1090
Db 644 SerThrThrGluSerSerAlaProValProValProValProValProValProValProValProVal 661
QY 1091 AACTCGCTACAGGCGACGCGTTCGCCCGC-----TCCAT 1126
Db 662 GluSerSerAlaProValProValProValProValProValProValProValProValProValProVal 681
QY 1127 TCCGCTACGTCACAAATACCGAAACCGCTGCTCAAAACCGCGCATCAAACTCAGC----- 1181
Db 681 aProValThrSerSerThrThrGluSerSerAlaProValThrSerSerThrThrThrThrThrThrThr 701
QY 1182 -AGTGCAACCAACCACTCCGTCGCCGCCCAAAAGACAAAGACGCTGACCGGAGCTCC 1240
Db 701 uSerSerAlaProValProValProValProValProValProValProValProValProValProVal 721
QY 1241 GTACACGCTGCATGTCACAAACCTTCATTTGTGACGCGCAACGCATCTTCATCGGTCA 1300
Db 721 o-----ValProThrProSer-----SerSerThrThrThrThrThrThrThrThrThrThr 731
QY 1301 TTCAACCTCGACCCCGTTCCGACGCGCTCAATACCGAAATGGCGGTGTCATCGCAAGC 1360
Db 731 uSerSerSerAlaProValProValProValProValProValProValProValProValProValProVal 751
QY 1361 CCCAAAATCGCAGAACAGATGGAGCGCACCTTCGCCGATACCCACCCCGAATAGC----- 1415
```

```
Db 751 oValThrSerSerThrThrGluSerSerAlaProValProValProValProValProValProValThr 771
QY 1416 -----CCTACCGGTTTACCTCGACAAACAACACC-----GCCTGCAA 1453
Db 771 rThrGluSerSerAlaProValProValProValProValProValProValProValProValProVal 791
QY 1454 TGGCACGATCCCGCCACCCGAAAAACCTACCGCAACGAAC-----CCGAA 1498
Db 791 xAlaProValProValProValProValProValProValProValProValProValProValProVal 811
QY 1499 GCCA-----AACTTTGAAACGATCCCGCAAAAATCC 1532
Db 811 rProSerSerSerSerSerSerSerSerSerSerSerSerSerSerSerSerSerSerSerThr 826
RESULT 15
G70063
cardiolipin synthase homolog ywnE - Bacillus subtilis
C:Species: Bacillus subtilis
C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 09-Jul-2004
C:Accession: G70063
R:Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Bertolotti, C.; Bron, S.; Brouillette, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; C. A.; Ehrlich, S.D.; Emmerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrati, E. Nature 390, 249-256, 1997
A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gallucci, J.; Harwood, C.R.; Henaout, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.; Koetter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinot, Y. M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelli, Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanlon, A:Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowaka, A.; Serca, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama, T.; Winters, P.; Wipet, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, A:Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.
A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis
A:Reference number: A69580; MUID:98044033; PMID:9384377
A:Accession: G70063
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-482 <RUN>
A:Cross-references: UNIPROT:P71040; UNIPARC:UPI0000060B8F; GB:Z99122; GB:AL009126; NID
A:Experimental source: strain 168
C:Genetics:
A:Gene: ywnE
C:Superfamily: cardiolipin synthase
Alignment Scores:
Pred. No.: 5,78e-11 Length: 482
Score: 314.00 Matches: 126
Percent Similarity: 38.4% Conservative: 65
Best Local Similarity: 25.3% Mismatches: 189
Query Match: 11.0% Indels: 118
DB: 2 Gaps: 15
US-10-665-990A-13 (1-1561) x G70063 (1-482)
QY 113 GAACGACGAAAGCGTCATTTCAATCTTCCAAACCTCTCTCTCGGACCAACATCTG 172
Db 83 GluAspLeuGluThrLysGlnPheGlnPheAsnAsnArgAlaThrPheAspAsnLysAsp 102
QY 173 CAATCCGCGCACACCCCTCATATAACACGGGCTA-----TCGACATCTAC 217
Db 103 LeuIleTyMetLeuIleMetAsnAsnHisAlaValPheThrGluAspAsnSerValAsp 122
QY 218 CTGCTCGACGACCCCAAGAGCCCTTGGCCGCCCGCGCCCTTATTCGAATCTGCGCAA 277
Db 123 ValIleThrAspGlyArgAspLysPheGlnArgLeuLeuSerAspIleSerLysAlaLys 142
QY 278 CACAGCTCGATTTGCATATCTACATTTGGCGCAACGACATTTCCGCGCAGCTGCTTC 337
Db 143 AspHisIleHisLeuGlnTyTrIleTyLysGlyAspLeuLeuGlyLysLeuArg 162
QY 338 AACCTCATGTACCTTGGCGGAAACGCGGTGCGGTACGCTGCTGTTGGACGACAC 397
```



```
QY 590 AATATCGCGCAGATACCTTCAAGTCGGTGAGGACACCGCTTTTCGCC-----GAC 640
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
D 251 AsnValGlyAspGluTyr-----LeuGlyLysSerLysPheGlyTyrTrpArgAsp 268
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 641 CTGACATCTCCGCGCAGCGGTCGTCGCGGAGTATCGCAGGATTCGACGGCTAC 700
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
D 269 ThrHisLeuArgGileValGlyAspAlaValAsnAlaLeuGlnLeuArgPheIleLeuAsp 288
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 701 TGGGCAAGCCATTCGCGCCACACGCGCAGCGC----- 733
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
D 289 TrpAsnSer-----GlnAlaThrArgAspHisIleSerTyrAspArgTyr 304
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 734 -----ATCATCCGAGCGGCAACATCGCGAGGCTTCAAGCAGCTCGATACAAC 784
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
D 305 PheProAspValAsnSerGlyThrIleGlyValGlnIleAlaSerSerGlyProAsp 324
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 785 GACGAAACATCGACACCGCTCTGCGCTACCGGAAACCGTCGAAAGTCGCCCTC 844
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
D 325 GluGlu----- 326
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 845 TACCAAAAATACAGACGGAGCGATCGACTGGCAGAGCGTCCAAACCGCTCATCAGC 904
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
D 327 TrpGluGlnIleLysTyrGlyTyrLeu-----LysMetIleSer 339
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 905 GACACCCCTGCAAAAGGACTCGACCGCGACCGCGCAACCGCGGAGGCTG 964
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
D 340 SerAla----- 341
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 965 CAAGACGGCTCAACAGCCCGAAAAAGCTCTATCTGTTTCAACCTATTTCGTCCT 1024
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
D 342 -----LysLysSerIleTyrIleGlnSerProTyrPheIlePro 354
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 1025 ACAAAATCGGCACAGACGACTCGCAAAACTGTGTGAGGACGATAGCTTACCGTC 1084
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
D 355 AspGlnAlaPheLeuAspSerIleLysIleAlaLeuGlyGlyValAspValAsnIle 374
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 1085 CTGACCAACTCGCTACAGCGACCGAGCTTCCCGCGCTCCATTCGCGTACGCAATAC 1144
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
D 375 MetIleProAsn-----LysProAspHisProPheValPheTrpAlaThrLeuLysAsn 392
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 1145 CGAAACCGCTGTCAAGCGGATCAAACTACGAGTGTGCAACCCACCATGCCGCT 1204
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
D 393 AlaAlaSerLeuLeuAspAlaGlyValLysValPheHisTyr----- 406
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 1205 CCGCCCAAAAGACAAAGGCTGACCGGAGCTCCGTAAACCGCTGATGCCAAAC 1264
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
D 407 -----AspAsnGlyPhe-----LeuHisSerLysThr 415
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 1265 TTCATTGTGAGCGCAACGATCTTCATCGGCTCATTCACCTCGACCCCGCTCGCA 1324
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
D 416 LeuValIleAspAspGluIleAlaSerValGlyThrAlaAsnMetAspHisArgSerPhe 435
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 1325 CGGCTCAATACCGAAATGGGCGTGTATCGAAGAGCCCAAAATCGCAGACAGATGAG 1384
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
D 436 ThrLeuAsnPheGluValAsnAlaPheIleTyrAspGlnIleAlaLysLysLeuLys 455
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 1385 CGCACCTCGCGGAT-----ACCACCGGATACGCTAC 1420
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
D 456 GlnAlaPheIleAspLeuAlaValSerSerGluLeuThrLysAlaArgTyrAlaLys 475
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 1421 CGC 1423
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
D 476 Arg 476
```

RESULT 17

cardiolipin synthase homolog lin0008 [imported] - Listeria innocua (strain Clip11262).
C:Species: Listeria innocua
C:Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 09-Jul-2004
C:Accession: A11433
R:Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.

D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A:Authors: Krest J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; P
ok, C.; Schluter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland
A:Title: Comparative genomics of *Listeria* species.
A:Reference number: AB1077; MUID:21537279; PMID:11679669
A:Accession: A11433
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-504 <GLA>
A:Cross-references: UNIPROT:Q92FUS; UNIPARC:UPI00000CCOB3; GB:AL592022; PIDN:CAC95241.1
A:Experimental source: strain Clip11262
C:Genetics:
C:Gene: lin0008
C:Superfamily: cardiolipin synthase

Alignment Scores:
Pred. No.: 3.84e-10 Length: 504
Score: 300.00 Matches: 118
Percent Similarity: 37.9% Conservative: 76
Best Local Similarity: 23.0% Mismatches: 210
Query Match: 10.5% Indels: 108
DB: 2 Gaps: 16

US-10-665-990A-13 (1-1561) x A11433 (1-504)

```
QY 26 ACTATATCCCGATGAAACACACGAGCTCATTTCCCTTTTATGCTCTCTCTGTTC 85
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
D 90 ThrAlaGlnValMetGluLysAlaLysLeuIleAsnAlaIleHisAlaIleProAsn 109
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 86 TGTTCCTCATGTTGCCCCCACTGGAAACGAGCGGAAACCGCTCATTTCAATATTC 145
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
D 110 ThrAsnGluLysLeuProArgLeuSerLysArgIle-----AlaHisLeuThr 127
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 146 AATCCTGCTCTCTGGACAAACATCTCGCAAAATCGGCACACCCCTATAACAAC 205
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
D 128 GluProIle-----LysGly--- 132
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 206 TCCGACATCTACCTGCTCGACGACCCCGACGAGCCCTTCGCCCGCGCGCTTATC 265
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
D 133 AsnLysIleGluIleLeuThrAsnGlyGluGluThrPheProValLeuLeuAspAla 152
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 266 GAATCTGCGCAACACACGCTCGATTTCGCAATCTACATTTGGCGCAACGACAT 325
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
D 153 ArgLysAlaGluAsnHisIleHisIleGlnTyrTyrIlePheLysThrAspAla 172
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 326 AGGCTGCTGTAACTCATGTACCTTTCGCGAGAACGCGGTGCGGTACGCTGCTG 385
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
D 173 ThrGluIleArgAspIleLeuValGluLysAlaLysSerGlyValGluValArg 192
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 386 TTGACGACAAACACGCGCGGTGGACGATCTCTGCTCGCTCGCTCGACGCCATCCC 445
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
D 193 PheAspGlyLeuGlySerLysLeuSerLysAlaPheLeuAlaProLeuLysGlu 212
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 446 AATATCGAGTCGCGCTGTTCACCCCTTCGCTCTACGCAAAATGGCGCGCATC 505
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
D 213 GlyValSerIleHisAlaPheAspPro---IleAlaSerProTrpIleVal----- 228
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 506 CTGACGACTTCCCGCGCTCAACCGCGCATCGACAAACAAATCCTTTACCGCGCA 565
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
D 229 -----ArgThrAlaAsnLeuArgAsnHisArgLysIleValValIleAsp 244
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 566 CGCGCCACCATCTCGCGGACGCAATATCGGCGCGCAATAC----- 607
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
D 245 GlnIleGlyPheThrGlyGlyLeuAsnIleGlyGluGluTyrArgSerAsnThrPro 264
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 608 TTCAAATCGGTCGAGCACACCGTTTTCGCGCATCTCGCATCTCGCCACCGCGCGCTC 667
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
D 265 PheArgValTrpArgAspThr-----HisMetLysIleThrGlyGlnAla 279
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 668 GTCGCGGAAGTATCGCACGACTTC-----GACCGTACTGGCAAGCCATTCGCGCC 721
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
D 280 ValIleGluLeuGlnGluSerPheLeuAsnAspTrpIleTyrMetGluAsnGlnAla 299
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
```

[illegible][illegible]

```
Db 242 ThrSerValGlnProSerSerSerSerSerAlaProThrThrThrSerAlaThrSerValGln 261
QY 825 CGTGAACAGTCGC-----CCCTTACCAGAAATACAGACGGGACGCATCG 872
Db 262 ProSerSerSerSerThrProThrProThrThrSerValGlnProSerSerSer 281
QY 873 ACTGGCAGAGCGTCCAAACCCGCTGATCAGCGACACCCCTGCAAAAGGACTCGACCGCG 932
Db 282 -----SerSerAlaProThrThrSerAlaThrSerValGlnProSerSerSer--- 297
QY 933 ACCGCGCAAAACCGCGGATTGCGGGAGGCTGCAAGACGCGCTCAACAGCCGCAAAAAA 992
Db 298 -----SerSerProThrProThrSerSerProThrProThrSer 303
QY 993 GCGTCTATCTGTTTACCTTATTCGTCCTTACAAATCCGCGACACGCACTGGCAA 1052
Db 304 SerThrIleSerValGlnProSerSerSerSerSerProThrThrSer----- 320
QY 1053 AACTGGTCAGGACGGCATAGACGTTACCGTCTGACCAACTCGCTACAGGCGGACCGACG 1112
Db 321 -----ThrThrSerValGlnProSerSer 327
QY 1113 TTGCGCGCGTCAATTCGCGCTACGTCAAATACC----- 1145
Db 328 SerSerGlySerAlaProThrThrSerAlaThrSerValGlnProSerSerSerSer 347
QY 1146 -----GAAACCCGCTGCTCAAGCCGCGATCA 1172
Db 348 ProProIleSerSerThrIleSerValGlnProSerSerSerSerSerProThrThr 367
QY 1173 AACTCTACGAGCTGCAACCCCAACCATGCGTCCCGCCACAAAAGACAAAGCGCTGACCG 1232
Db 368 SerThrThrSerValGlnProSerSerSerGlySerAlaProThrThrSerAlaThrSer 387
QY 1233 GCAGCTCGTAAACGAGCTGATGCGCAAAACCTTCAATGTGGAGGCGCAAAAGCATCTCA 1292
Db 388 ValGlnPro-----SerSerSerSerSerValProThrThrThrSerAlaThrSer 403
QY 1293 -----TCGGTCTCATCACTCGACCCCGCTTCGCGACGCTCAATACCGAATGGCG 1346
Db 404 ValArgSerSerSerSerSerSerSerProThrThrThrThrSerValGlnProSer 423
QY 1347 TCGTCATGAAAGCCCAAAATCG-----CAGAACAGATGAGCGGCACCC 1391
Db 424 SerSerSerSerValProThrThrSerAlaThrSerValGlnThrSerSerSerSer 443
QY 1392 TCGCGGATACCA-----CACCGAATACGCTACCGGTTACCTCGACCA 1436
Db 444 ThrProIleProSerThrThrSerValGlnProSerSerSerSerSerAlaProThrThr 463
QY 1437 AACACAACCGCTCG-----AATGGCAGATCCCGCCACCGCAAAACCTACCGAAGC 1490
Db 464 SerAlaThrSerValGlnProSerSerSerSerSerProIleSerSerThr---Ile 482
QY 1491 AACCCGAAGCCAAACTTTGGAACGCAATCGCGCAAAATCC 1532
Db 483 SerValGlnProSerSerSerSerSerSerProThrThrSer 496
```

RESULT 19

T45462
membrane glycoprotein [imported] - equine herpesvirus 1
C:Species: equine herpesvirus 1
C:Date: 31-Jan-2000 #sequence_revision 31-Jan-2000 #text_change 09-Jul-2004
C:Accession: T45462
R:Kirisawa, R.; Kobayashi, T.; Kawakami, Y.; Iwai, H.
J. Equine Sci. 7, 79-87, 1996
A:Title: Nucleotide sequences of open reading frames 1, 24 and 71 of an attenuated equine herpesvirus 1
A:Reference number: 222973
A:Accession: T45462
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-866 <KIR>

A:Cross-references: UNIPROT:O39781; UNIPARC:UPI00000ECBA1; EMBL:D88733; PIDN:BAA20037.1
A:Experimental source: strain HH1
C:Genetics:
A:Note: ORF71
C:Superfamily: equine herpesvirus glycoprotein X; equine herpesvirus 1 glycoprotein homology
F:558-866/Domain: equine herpesvirus 1 glycoprotein homology <EHG>

Alignment Scores:

Pred. No.:	5.6e-10	Length:	866
Score:	297.50	Matches:	122
Percent Similarity:	39.2%	Conservative:	72
Best Local Similarity:	24.6%	Mismatches:	222
Query Match:	10.4%	Indels:	79
DB:	2	Gaps:	12

US-10-665-990A-13 (1-1561) x T45462 (1-866)

```
QY 126 GCCGTCAATTCATPACTTCCAAACCTGTCTCTCGACAAACATCTGCAATTCGGCACA 185
Db 19 AlaileGlySerThrThrThrThrThrGluThrThrThrSerSerSerSerThrSerGlySer 38
QY 186 CCCCTATAACAACGGGCTATCCGACATCTACCTGCTCGACGACCCCAAGACCCCTTG 245
Db 39 GlyGlnSerThrSerSerGlyThrThrAsnSerSerSerSerProThrThrSerPro--- 57
QY 246 CCGCCCGCGCGCCCTTATCGAATCTCGGAACACAGCCCTCGATTTCGAATACTACATT 305
Db 58 -----ProThrThrSerSerSerProProThrThrThrHisThrSerSerProSer 74
QY 306 GCGCAACAGACATTTTCGCGCAGGCTGTGTTCAACCTCATGTACTTCCCGCAGAACGCG 365
Db 75 SerThrSerThrGlnSerSerSerThrAlaAlaThrSerSerSerSerAlaProSerThrAla 94
QY 366 GCGTCCGCTACGCGCTCTGTGACAGCAACACGCGGGGTGGACGATCTCCTGC 425
Db 95 -----SerSerThrSerSerSerSerSerSerSerSerSerSerSerSer 106
QY 426 TCGCCCTCGACGACATCCCAATATCGAAGTGGCGCTTCAACCCCTCGTCTCTAGCA 485
Db 107 GluThrThrThrThrThrProThrAlaSerThrThrThrProThrThrThrThrAlaAla 126
QY 486 AATGGCGCGCACTCGGCTACCTGACGACTTCCCGCCTCAACCGCGCATGCAACA 545
Db 127 -----ProThrThrAlaAlaThrThrThrThrAlaValThrThr 138
QY 546 AATCCTTTACGCGGACACCGGCCACCATATCTCGCGGACGCAATATCGGACCAAT 605
Db 139 -----AlaAlaSerThrAlaAlaSerThrSerAlaGluThrThrThrAlaThrAla 155
QY 606 ACTTCAAAGTCGGTGAGGACACCGTTTTCGCCGACCTGGACATCTCTCGCCACCGCAGCG 665
Db 156 ThrAlaThrSerThrProThrThrThrThrProThrSerThrThrThrThrThrThrAlaThr 175
QY 666 TCGTCGGCGAAGTATCGCAGCTTCGACGCTACTGGGCAAGCCATTCGCGCCCAACG 725
Db 176 ThrThrValPro---ThrThrAlaSerThrThrThrThrThrThrThrThrThrThrThr 194
QY 726 CCAGCGCATATCCGAGCGGCAACATCCGGAAGGGTCTTCAAGCATCGGATACAAACG 785
Db 195 ThrAlaAlaThrThrThrAlaAlaThrThrThrThrAlaAlaThrThrThrThrThrThr 214
QY 786 ACAGAAACATCCAGACGCGCTCTCGCTACCGCGGAAACCGTCGAACAGTCGCCCTCT 845
Db 215 ThrAlaAlaThrThrThrAlaAlaThrThrThrThrAlaAlaThrThrThrThrThrThr 234
QY 846 ACCAAAAATACAGACGGGACGATCGATCGGAGAGCGTCCAAACCCGCTCATCATCAGCG 905
Db 235 ThrAlaAlaThrThrThrAlaAlaThrThrThrThrSerSerAlaThrThrThrThrThr 254
QY 906 ACACCCCTGCAAAAGACTCGACCGGACCGCGCGCAACCCCGGATTGCGGGAGGCTGC 965
Db 255 AlaAlaThrThrThrAlaAlaThrThrThrThrAlaAlaThrThrThrThrThrThrThr 271
```



```
QY 1277 GCGAAGCGACTTTCATCGGCTCATTAACCTCGACCCCGCTTCGCGAGCGCTCAATACC 1336
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 429 GlyThrArgAlaAlaIleGlyThrAlaAsnPheAspValArgSerPheArgLeuAsnHis 448
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1337 GAAATGGCGGCTCGTCAAGAGCCCAAAATCGCAGAACACAGATGGAGCGCACCCCTGCC 1396
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 449 GluLeuMetValPheLeuTyArgGlu-----SerGluAlaMetHisLeuLysArg 466
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1397 GATACCAACCGCAATACCGCTACCGCTTACCCTTCGACAAACACACACCGCTGCAATGG 1456
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 467 AspPheLysLysAspPhe-----GluAspSerArgLeu----- 477
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1457 CACGATCCGCGCACCGGAAACCTACCGACGAAACCGAAGCCAAACTTGGAAACGC 1516
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 478 -----PheThrMetLysAspMetGluAsnLysProLeuLeuThrArgIleLysGlu 494
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1517 ATCGCGCGCAAAATCCTATCC 1537
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 495 ValLeuSerSerLeuLeuSer 501
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 22
A11762
cardiolipin synthase homolog lin2646 [imported] - Listeria innocua (strain Clip11262)
C:Species: Listeria innocua
C:Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 09-Jul-2004
C:Accession: A11762
R:Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker,
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.
D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A:Authors: Krefz, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Ma
ok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland,
A:Title: Comparative Genomics of Listeria species.
A:Reference number: AB1077; MUID:21537279; PMID:11679669
A:Accession: A11762
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-482 <GLA>
A:Cross-references: UNIPROT:Q92720; UNIPARC:UPI00000CC933; GB:AL592022; PIDN:CAC97873.1;
A:Experimental source: strain Clip11262
C:Genetics:
A:Gene: lin2646
C:Superfamily: cardiolipin synthase
```

```
Alignment Scores:
Pred. No.: 1.13e-09 Length: 482
Score: 292.00 Matches: 109
Percent Similarity: 39.0% Conservative: 56
Best Local Similarity: 25.8% Mismatches: 164
Query Match: 10.2% Indels: 94
DB: Gaps: 15
```

US-10-665-990A-13 (1-1561) x A11762 (1-482)

```
QY 206 TCCGACATCTACTGCTCGACGACCCCGACGAGCCCTTGGCGCGCGCCCTTATC 265
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 119 AsnGluValGluPheIleAspGlyHisGluLysPheAspAlaLeuIleAlaAspIle 138
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 266 GAATCTGCGCAACACAGCTCGATTTGCAATACTACTTTGCGCAACGACATTTCCGGC 325
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 139 GluLysAlaLysAspHisIleHisLeuIleTyTyIlePheHisSerAspGluLeuGly 158
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 326 AGGCTGTGTTCACCTCACTACCTTGGCGGAGACCGCGGTGGGTAGCCTGTGCTG 385
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 159 AsnArgLeuMetArgValLeuGluArgLysAlaAlaGluGlyLeuAsnValLysIleIle 178
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 386 TTGGACGACAAACACGCGCGGTGGACGATCTCTCTCGCCCTCGACGACCATCCC 445
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 179 TyrAspAlaMetGlySerArgThrThrLysLysSerPhePheArgThrPheGluLysAsn 198
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 446 AATATCGAAGTGGCGGTGTTCAACCCCTTCGTCTACGCAAAATGGCGGCACTCGGCTAC 505
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 199 GlyGlyLeuValArgProPhePheProSerLysLeuPro----- 211
```

RESULT 23

AG1387

cardiolipin synthase homolog lmo2503 [imported] - Listeria monocytogenes (strain EGD-e)
C:Species: Listeria monocytogenes

```
QY 506 CTGACCGACTTCCCGCCGCTCAACCGCGCATGCACAAACAATCCTTTACCGCGCAGAAC 565
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 212 LeuIleAsnPhe---ArgLeuAsnTyArgAsnHisArgLysLeuAlaIleIleAspGly 230
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 566 CGCGCCACCATACATCGCGGACGCAATATATCGGCACGAATACTTTCAAAGTCGGTGAGGAC 625
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 231 AspValGlyTyriIleGlyGlyPheAsnIleGlyAspGluTy-----LeuGlyArgSer 248
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 626 ACCGTTTTTCGCGCACCTGGACATCTCGCCACCGCGACGCTGCTGGCGAAGTATCGCAC 685
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 249 LysLysPheGly----- 252
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 586 GACTTCGACCGCTACTGGGCAAGCCATTCCGCCACCAACGCCACCGCGCATCATCGCAGC 745
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 253 -----TyrrTpArgAsp-----ThrHisLeuArgValHis 262
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 746 GGCAACATCGGCAAGGCTCTTCAAGCA-----CTCGGATACAAACGACGCAACATCC 796
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 263 GlyLysAlaValTyAlaMetGlnThrArgPheIleMetAspTrpAsnSerAlaSerSer 282
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 797 AGACACGCGCTC-----CTCGCTACCGCGAAACCGTCTC-----GAACAGTCGCGCC 841
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 283 ThrHisLysIleAspTyLysAlaArgTyPheProThrPheHisGlyLysGlyHisThr 302
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 842 CTCTACCAAAAATACAGACGGGA---CGCATCGACTGGCAGAGGTCACAAACCGCGCTG 898
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 303 SerMetGlnIleValSerSerGlyProAspSerGluTrpGlnGlnIleLysAsnGlyTy 322
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 899 ATCAGCGACACCCCTCGAAAAGGACTCGACCGCGACCGCGCAAAACCGCGGATTCGCGG 958
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 323 IleLys----- 324
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 959 AGGCTGCAAGACGCGCTCAAAACAGCCCGGAAAGCGTCTATCTGTTTCACTATTTC 1018
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 325 -----MetIleAsnAlaLysLysThrIleTyLeuGlnSerProTyPhe 340
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1019 GTCCCTACAAATCGGCACAGACGCTGCAAAACTGGTGCAGGACGATGACGCTT 1078
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 341 IleProAspAlaSerLeuLeuGluAlaIleLysIleAlaAlaLeuSerGlyValAspVal 360
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1079 ACCGTCCTGACCAACTCGCTACAGCGACCGACGCTGCGCGCTCCATTCGCGGTACGTC 1138
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 361 ArgValMetIleProAsn-----LysProAspHisAlaPheValTyArgAlaThrThr 378
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1139 AAATACGAAACCGCTGCTCTCAAGCGCGCATCAAACTCTACGAGCTCCAAACCAACCAT 1198
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 379 AsnTyAlaGlyGluLeuMetGluThrGlyAlaLysIlePhe----- 392
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1199 GCGTCCCGCGCACAAAGACAAAGCGCTGACCGCGAGCTCCGTAAACAGCTGCATGCC 1258
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 393 -----IleTyAspAsnGlyPhe-----IleHisAla 401
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1259 AAAACCTTCAATGTGGACGCAAAACCGCATCTTTCATCGGCTCATTCACCTCGACCCCGT 1318
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 402 LysThrLeuValValAspGlyGluIleAlaSerValGlyThrAlaAsnMetAspPheArg 421
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1319 TCCGACCGCTCAATACCGAAATGGCGTCTCGAAAGCCCAAAATCGCAGAACAG 1378
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 422 SerPheArgLeuAsnPheGluValAsnAlaPheIleTyGluLysLysMetValGlnLys 441
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1379 ATGGAGCGCACCTCGCCGAT-----ACCACACCCGAA---TAC 1414
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 442 LeuGluAspAlaPheLeuGluAspIleLeuLysSerTyGlnLeuThrProGluLeuTy 461
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1415 GCCTACCGC 1423
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 462 AlaLysArg 464
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

C>Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 09-Jul-2004
C/Accession: AG1387

R;Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecker, D.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A/Authors: Kretz, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Makok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland, A.; Title: Comparative genomics of *Listeria* species.
A/Reference number: AB1077; MUID:21537279; PMID:11679669

A/Accession: AG1387

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-482 <GLA>

A/Cross-references: UNIPROT:Q8Y4E3; UNIPARC:UPI000055715; GB:NC_003210; PIDN:CAD00581.1

A/Experimental source: strain EGD-e

C/Genetics:

A/Gene: lmo2503

C/Superfamily: cardiolipin synthase

Alignment Scores:

Pred. No.:	2,21e-09	Length:	482
Score:	287.00	Matches:	108
Percent Similarity:	39.0%	Conservative:	57
Best Local Similarity:	25.5%	Mismatches:	164
Query Match:	10.1%	Indels:	94
DB:	2	Gaps:	15

US-10-665-990A-13 (1-1561) x AG1387 (1-482)

Qy	206	TCCGACATCTACTGCTCGACGACCCCAAGCCCTTGGCCGCGCGCCCTTATC	265
Db	119	AsnGluValGluLeuPheValAspGlyHisGluLysPheAspAlaLeuAlaAspIle	138
Qy	266	GAATCTCGGACACAGCCTCGATTGCAATACATATTGGCCGACGACATTTCGGC	325
Db	139	GluLysAlaLysAspHisHileHisLeuIleTyrIlePheHisSerAspGluLeuGly	158
Qy	326	AGGCTGTGTTCAACCTCATCTACCTTGGCCGACGCGCGGTGCGGTACGCTGCTG	385
Db	159	AsnArgLeuMetArgValLeuGluArgLysAlaAlaGluGlyLeuAsnValLysIle	178
Qy	386	TTGACGACAAACACACGCGGGTTGGACGATCTCTGCTCGCCCTCGACGACATCCC	445
Db	179	TyrAspAlaMetGlySerArgThrThrLysLysSerPheArgThrPheGlnLysAsn	198
Qy	446	AATATCGAAGTGGCCTGTCAACCCCTTCTGCTCTACGCAATGGCGGCACTCGGTAC	505
Db	199	GlyGlyLeuValArgProPheProSerLysLeuPro	211
Qy	506	CTGACCGACTTCCCGCGCTCAACCGCGCATGCAACAATCTTTACCGCGCAAC	565
Db	212	LeuIleAsnPhe--ArgLeuAsnTyrArgAsnHisArgLysLeuAlaIleAspGly	230
Qy	566	CGCGCCACCATCTCGCGGACGCAATATCGGACGCAATCTTCAAGTGGTGAGGAC	625
Db	231	AspValGlyTyrIleGlyGlyPheAsnIleGlyAspGluTyr-----LeuGlyAlaSer	248
Qy	626	ACCGTTTTCGCGACCTGGACATCTCGCCACCGCGGCGCTGTCGGCGAAGTATCCAC	685
Db	249	LysLysPheGly-----	252
Qy	686	GACTTCGACCGCTACTGGGCAAGCATTCGCCCCCAACGCGCATCATCCGACG	745
Db	253	-----TyrTrpArgAsp-----ThrHisLeuArgValHis	262
Qy	746	GGCAACATCGCAAGGCTTCAAGCA-----CTCGGATACACGACGAAACATCC	796
Db	263	GlyLysAlaValTyrAlaMetGlnThrArgPheIleMetAspTrpAsnSerAlaSerSer	282
Qy	797	AGACACGGCTC-----CTCGGCTACCGCAACCGCTC-----GAACAGTCCGCC	841
Db	283	ThrHisLysIleAspTyrLysAlaArgTyrPheProThrPheHisGlyLysGlyHisThr	302

Qy	842	CTCTACCAAAAAATACAGACGGGA---CGCATCGACTGGCAGAGGGTCCAAACCCGCTG	898
Db	303	SerMetGlnIleValSerSerGlyProAspSerGlnGlnGlnIleLysAsnGlyTyr	322
Qy	899	ATCAGCAGACACCCCTGCAAAAGGACTCGACCGCAGCCGCAAAACCGCGATTGCCGGG	958
Db	323	IleLys-----	324
Qy	959	AGGCTGCAAGACGCGCTCAACACGCCCGAAAAAAGCGTCTATCTGGTTTCAACCTATTTC	1018
Db	325	-----MetIleAsnAlaIleLysLysThrIleTyrLeuGlnSerProTyrPhe	340
Qy	1019	GTCCTCAAAAAATCCGCGACAGCGCTGCAAAACTGGTGCAGGACGGCATAGACCTT	1078
Db	341	IleProAspAlaSerLeuLeuGluAlaIleLysIleAlaLeuSerGlyValAspVal	360
Qy	1079	ACGCTCTGACCAACTCGCTACAGGCGACCGACGCTTGCCTGCTCCATTCCGGCTACGTC	1138
Db	361	ArgValMetIleProAsn-----LysProAspHisAlaPheValTyrArgAlaThrThr	378
Qy	1139	AAATACCGAAACCGCTGCTCAAGCGCGCATCAAACTCTACGAGCTGCAACCCCAACCAT	1198
Db	379	AsnTyrAlaGlyGluLeuMetGluThrGlyAlaLysIlePhe-----	392
Qy	1199	GCGTCCCGCCCAACAAAGAGCGCTGACCGGCGAGTCCGTACACGAGCTCGCATGCC	1258
Db	393	-----IleTyrAspAsnGlyPhe-----IleHisAla	401
Qy	1259	AAAACTTCATTGTGGACGGCAACGCATCTTCATCGGCTCATTCACCTCGACCCCGCT	1318
Db	402	LysThrLeuValValAspGlyGluIleAlaSerValGlyThrAlaAsnMetAspPheArg	421
Qy	1319	TCCGACGGCTCAATACCGAATGGGCGTGTCTATCGAAAGCCCAAAATCGCAGACAG	1378
Db	422	SerPheArgLeuAsnPheGluValAsnAlaPheIleTyrGluLysGlnMetValGlnLys	441
Qy	1379	ATGAGCGCGACCCCTCGCGCAT-----ACCACACCGGAA---TAC	1414
Db	442	LeuGluAspAlaPheLeuGluAspIleLeuLysSerTyrGlnLeuThrProGluLeuTyr	461
Qy	1415	GCCTACCGC	1423
Db	462	AlaLysArg	464

RESULT 24
T45025
mucin MUC5B, tracheobronchial [imported] - human (fragment)
C/Species: Homo sapiens (man)
C/Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 21-Jul-2000
C/Accession: T45025
R;Desseyn, J.L.; Guyonnet-Duperat, V.; Porchet, N.; Aubert, J.P.; Laine, A.
J. Biol. Chem. 272, 3168-3178, 1997
A/Title: Human mucin gene MUC5B, the 10.7 kb large central exon encodes various alternative splicing products
A/Reference number: 222899; MUID:97166151; PMID:9013550
A/Accession: T45025
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-3570 <DES>
A/Cross-references: UNIPARC:UPI000016AD88; EMBL:Z72496; NID:g1834502; PIDN:CAA96577.1;
A/Experimental source: placenta
C/Genetics:
A/Gene: MUC5B

Alignment Scores:
Pred. No.: 2.94e-09 Length: 3570
Score: 286.00 Matches: 155
Percent Similarity: 31.0% Conservative: 67
Best Local Similarity: 21.7% Mismatches: 211
Query Match: 10.0% Indels: 282
DB: 2 Gaps: 27

US-10-665-990A-13 (1-1561) x T45025 (1-3570)

312	ValLysTyrPheProLysLysAsnGlyProLeuGly-	-----	323
782	AACGACGAAACATCAGACACGCGCTCTCGGCTACGCGAAACCGTCGAAACAGTCGCC	841	
324	-----	-----AsnSerPro	326
842	CTTACCAAAAAATACAGACGGGACGCATCGACTGCGAGAGGCTCAA-	-----ACC	892
327	IleGlnIleAlaLysSerGlyProLaserAspTyrHisGlnIleGluTyrGlyTyrThr		346
893	CGCTGATCGAGCACCCCTGCGAAAGGACTCGACCGCGACCGCGGCGAAACCGCGGATT	952	
347	LysMetIleMetSerAla-	-----	352
953	GCGCGGAGGCTGCAAGACGCGCTCAACACAGCCGAAAAAAGCGTCTATCTGTTGTTTCACCC	101	
353	-----	-----LysLysSerValTyrLeuGlnSerPro	361
1013	TATTTCGTCCCTACAAAATCGGCACACAGACGACCTGGCGAAACCTGGTCGAGGACCGCATAT	107	

362	Yrphellep	Asp	Asn	Ser	Y	I	L	E	A	S	N	A	L	E	S	I	L	E	A	L	A	L	S	S	E	R	G	L	y	A	l		381
1073	GAC	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	1102	
382	Asp	Val	His	Leu	Met	I	L	E	P	r	o	C	y	s																		399	
1133	TAC	G	T	C	A	A	A	T	A	C	C	G	A	A	A	C	C	G	T	C	T	C	T	C	T	C	T	C	T	C	T	C	1152
400	Thr	Phe	Ser	Asn	Ala	Ser	Asp	Leu	Ser	Ser	G	L	y	V	a	l	L	y	S	i	L	e	T	y									415

Db 416 -----ThrTyrGluAsnGlyPhe-----ile 422

QY 1253 CATGCGCAAAACCTTCATTGTGGACGGCAACGCATCTTTCATCGGCTCATTCACACCTCGAC 1311

Db 423 HisSerLysMetCysLeuIleAspAspGluLeuValSerValGlyThrAlaAsnMetCasp 442

QY 1313 CCGCGTTCCGACGGCTCAATACCGAAATGGCGCTCGTCATCGAAAGCCCCAAAATCGCA 1371

Db 443 PheArgSerPheGluLeuAsnPheGluValAsnAlaPheValTyrAspGluAsnLeuAla 462

QY 1373 GAACAGATGGAGCGCACCTTCGCGCATACACACCGCAATACGCCATACCGCTACCGGTTACCCCTC 1431

Db 463 LysAspLeu-----ArgValAlaTyrGluHisAspIle 473

QY 1433 GACAAACACAAACCGCTGCAATGCGACGATCCCGCCACCCGGAAAAAACCCTACCGCAGAA 1491

Db 474 ThrLysSerLysGlnLeu-----ThrLysGluSerTyrAlaAsnArg 487

QY 1493 CCGGACCGCAAAACTTTTGGAAACGCATCGCGCAAAAATCCCTATCC 1537

Db 488 ProLeuSerValLysPheLysGluSerLeuAlaLysLeuValSer 502

RESULT 26

H86744

cardiolipin synthase [imported] - Lactococcus lactis subsp. lactis (strain IL11403)

C;Species: Lactococcus lactis subsp. lactis

C;Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 09-Jul-2004

C;Accession: H86744

R;Bolotin, A.; Wincker, P.; Mauger, S.; Jaillon, O.; Malarne, K.; Weissenbach, J.;

Genome Res. 11, 731-753, 2001

A;Title: The complete genome sequence of the lactic acid bacterium Lactococcus

A;Reference number: A86625; MUID:21235186; PMID:11337471

A;Accession: H86744

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-481 <STO>

A;Cross-references: UNIPROT:Q9CGY2; UNIPARC:UPI000000C695C; GB:AE005176; PID:g11403

A;Experimental source: strain IL1403

C;Genetics:

C;Superfamily: cardiolipin synthase

C;Gene: cIsA

Alignment Scores:

Pred. No.: 4,06e-09 Length: 481
Score: 282.50 Matches: 104
Percent Similarity: 39.0% Conservativity: 76
Best Local Similarity: 22.5% Mismatches: 167
Query Match: 9.9% Indels: 115
DB: 2 Gaps: 15

US-10-665-990A-13 (1-1561) x H86744 (1-481)

```
QY 194 ACAACGGGCTATCCGACATCTACCTGCTCGACGACCCCAACGAGCCCTTCCGCCCGC 253
DB 118 AsnThrGly-----VallyIlePheAsnAspGlyArgGlnLysPheAsnAlaLeu 134
QY 254 GCGGCCCTTATGAAATCTCCGAACACAGCCTCGATTTCGAATCTACATTTGGCGCAAC 313
DB 135 IleAspAspIleAsnLysAlaLysHisHisValHisMetGluTyrTyrIlePheArgMet 154
QY 314 GACATTTCCGGAGGCTGCTGTTCACCTCATGTACCTTTCGCCGAGAACGCGCGTGGCG 373
DB 155 AspArgMetGlyHisGluIleTyrAspAlaLeuLeuAlaAlaLysArgGlyValGlu 174
QY 374 CTACGCTCTCTGGAC-----GACACACACGCGCGGTTGGACGATCTCCTG 424
DB 175 ValLysLeuLeuIleAspAlaTrpGlySerAsnLysThrLys----- 188
QY 425 CTCGCCCTCGACAGCCATCCCAATATCGAAGTGGCCTGTTCAACCCCTTCGCTACGC 484
DB 189 -----MetSerAsnPhelysGluLeuGln----- 196
QY 485 AAATGGCGCGACTCGGTACTCGTACCGGCTTC----- 517
DB 197 -----GluAlaGlyHisValAlaGlnPhePheProLeuIleMetProLeuValAsn 214
QY 518 CCGCGCTCAACGGCGCATGCACACAAATCTTTACCGCGCAACACGCGCCACCATA 577
DB 215 ProArgThrAsnTyrArgLeuHisArgLysIleValIleAspGlyIleIleGlyTyr 234
QY 578 CTCGGCGGACGAATATCCGGACCAATACTTCAAGTCGGTGCAGGAC---ACCGTTTTC 634
DB 235 ThrGlyGlyPheAsnValGlyAspGluTyrAlaSerIleThrLysLysPheGlyTyrTrp 254
QY 635 GCGACCTGGACATCTCCGACCGGACGCTGTCGGGGAAGTATCGCACTTCGAC 694
DB 255 ArgAspAsnHisLeuArgLeuThrGlyAspIleValTyrSerLeuGlnHisArgPheIle 274
QY 695 CGCTACTGGCAAGCATTCGCCCCACACGCGCATCATCCGACGCGCAACATC 754
DB 275 LeuAspTrp----- 277
QY 755 GCGAAGGGTCTTCAAGCACTCGGATACACGAGAAACATCCAGACACGCGCTCCTCGCG 814
DB 278 -----AsnSerGlnHis----- 281
QY 815 TRACCGGAAACCGTCAACAGTCGCCCTCTAC---CAAAAATACAGAGGACGCATC 871
DB 282 HisPheGluValThrGluGlyGluProTyrPheProAspSerIleValGluGlyHis--- 300
QY 872 GACTGCGAGGGTCCAAACCGCTGATCAGGACACCCCTGCAAAAGACTCGACGCG 931
DB 301 -----ValAlaThrGlnLeuValThrSerGlyPro----- 310
QY 932 GACCGCGCAAAACCGCGGATTCGGGGAGGCTGCAAGACGCGCTCAAAACAGCGCGGAAA 991
DB 311 AspGluAspLysGluGlnIleLysLeuThrTyrMetLysMetIleSerGlyAlaGluArg 330
QY 992 AGGCTATCTGTTTCAACCTATTTTCGTCCTTACAAAATCCGGACACAGCAGCTGGCA 1051
DB 331 GluIleIleIleIleIleIleIleIleIleIleIleIleIleIleIleIleIleIleIle 350
QY 1052 AAACCTGGTCAGGCGGCATAGACGTTACCGTCCTGACCACTCGCTACAGGCGCGAC 1111
```

```
DB 351 LeuAlaLeuLeuSerGlyValGlnValLysLeuLeuIleProAsn-----LysProAsp 368
QY 1112 GTTCCGCGCTCCATTCGCGGTACGTCAAATACGAAACCGCTGCTCAAAAGCGCGCATC 1171
DB 369 HisProLeuValTyrTrpAlaThrTyrPheHisAlaAlaAspLeuValLysTyrGlyAla 388
QY 1172 AAACCTCTAGAGTGCACCAACCATGCCCTCCCGCCACAAAAGACAAAGGCTGACC 1231
DB 389 LysValTyr-----ThrTyrGluAsnGlyPhe--- 397
QY 1232 GGCAGCTCCGTAAACAGCCTGCATGCCAAACCTTCATTGTGGAGCGCAACGCATCTTC 1291
DB 398 -----ValHisSerLysThrLeuIleIleAspGlyGluTyrAlaSer 411
QY 1292 ATCGGCTCATTTCAACCTCGACCCCGTTCGCACGGCTCAATACCGAAATGGCGTCTGC 1351
DB 412 ValGlySerAlaAsnLeuAspTyrArgSerLeuGlnLeuCysPheGluAlaAsnValVal 431
QY 1352 ATCGAAACCCCAAAATCGCAGACAGATGGAGCGCACCTCGCGATACACACCCGAA 1411
DB 432 IleTyrAspTyrAspIleSerGlnLysLeuArgAsn----- 443
QY 1412 TAGCGCTACCGGTTACCTCTGCACAAACACACACCGCTGCAATGCGACGATCCCGCCACC 1471
DB 444 -----AspPheMetLysAspLeuLysLeuSerArgProLeuThr 456
QY 1472 CGAAAAACCTACCGCAACGAAACCGCAAACTTTGGAAACGCATCCGCGCAAAATC 1531
DB 457 LeuGluArgTyrGluGluArgSerLysLeuValArgPheLysGluGlyLeuAlaArgLeu 476
QY 1532 CTATCC 1537
DB 477 IleAla 478
```

RESULT 27

B84007 cardiolipin synthetase B82858 [imported] - Bacillus halodurans (strain C-125)

C:Species: Bacillus halodurans

C>Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 09-Jul-2004

C:Accession: B84007

R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Hirai

Nucleic Acids Res. 28, 4317-4331, 2000

A>Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and

A:Reference number: A83650; MUID:20512582; PMID:11058132

A:Accession: B84007

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-503 <STO>

A:Cross-references: UNIPROT:Q9K824; UNIPARC:UPI0000127B9A; GB:AP001516; GB:BA000004; NII

A:Experimental source: strain C-125

C:Genetics:

C:Gene: B82858

C:Superfamily: cardiolipin synthase

Alignment Scores:

Pred. No.: 9.81e-09 Length: 503
Score: 276.00 Matches: 110
Percent Similarity: 37.2% Conservativity: 57
Best Local Similarity: 24.5% Mismatches: 188
Query Match: 9.7% Indels: 94
DB: 2 Gaps: 12

US-10-665-990A-13 (1-1561) x B84007 (1-503)

```
QY 167 ATCTGCAATTCGGGCACACCCCTCATAAACACGGGCTATCC-----GACATCTAC 217
DB 124 LeuPheArgLeuAlaHisArgLeuAlaAsnAsnProIleSerPheSerThrAsnThrLys 143
QY 218 CTGCTCGACGACCCCAACGAGCCCTTCGCCCGCGCGCCCTTATGAATCTCGCGCA 277
DB 144 ValLeuThrAspGlyLysGluThrPheAlaHisIleLysGlnAlaLeuArgMetAlaThr 163
QY 278 CACACCTCGATTTCGAATACTACATTGTCGCAACGACATTTCCGCGAGGCTGTGTTTC 337
```



```
Db      164 HishisileHisLeuGluTy-TyrileValArgAspGluileGlyGlnGluileLys 183
      |||  ...  |||...|||...|||  |||...|||  |||...  ...
Qy      338 AACCTCATGTACCTTGGCGGAGACGGCGTGGCGTACCGCTGCTTGGAGACAAAC 397
      |||...|||...|||...|||...|||...|||...|||...|||...|||...
Db      184 GluileLeuMetGlnLysGluileHisValArgPheLeuTyArgGlyVal 203
      |||...|||...|||...|||...|||...|||...|||...|||...|||...
Qy      398 AACACGGCGGTTGGAGATCTCTGCTCGCCCTCGACAGCCATCCCAATATCAAGTG 457
      |||...|||...|||...|||...|||...|||...|||...|||...|||...
Db      204 GlySerTrpLysLeuSerLysSerTyriLeGlnAspLeuLysGlnAlaGlyValGluile 223
      |||...|||...|||...|||...|||...|||...|||...|||...|||...
Qy      458 CGCTGTTCACCCCTTCTGCTCTACGCAATGGCGGCACCTCGCTACCTGACCGATTC 517
      |||...|||...|||...|||...|||...|||...|||...|||...|||...
Db      224 ValProPheAlaProValLysLeu----- 231
      |||...|||...|||...|||...|||...|||...|||...|||...|||...
Qy      518 CCC-----CGCTCAACCGCGCATGTCACACAAATCCTTTACCGCGGACAAAC 565
      |||...|||...|||...|||...|||...|||...|||...|||...|||...
Db      232 ProPheleAsnHisThrileAsnTyArgAsnHisArgLysileileValileAspGly 251
      |||...|||...|||...|||...|||...|||...|||...|||...|||...
Qy      566 CGCGCCACCATACTCGCGGAGCAATATCGCGACGAATACTTCAAAGTCGGTGAGGAC 625
      |||...|||...|||...|||...|||...|||...|||...|||...|||...
Db      252 ThrValGlyPheValGlyGlyLeuAsnileGlyAspGluTy-----LeuGlyLysAsp 269
      |||...|||...|||...|||...|||...|||...|||...|||...|||...
Qy      626 ACCGTTTTCGCC-----GACTGGACATCTCGCCACCGCGAGCGTCTCGGCGAA 676
      |||...|||...|||...|||...|||...|||...|||...|||...|||...
Db      270 ProTyrrPheGlyPheTrpArgAspThrHisLeuTyrrValArgGlyGluAlaValArgThr 289
      |||...|||...|||...|||...|||...|||...|||...|||...|||...
Qy      677 GTATCGCAGCATTCGACCGCTACTGGCAAGCATTCGCGCCCAACAGCCACCGCATC 736
      |||...|||...|||...|||...|||...|||...|||...|||...|||...
Db      290 LeuGlnLeuilePheLeuArgAspTrpAlaHisGluThrGly----- 303
      |||...|||...|||...|||...|||...|||...|||...|||...|||...
Qy      737 ATCCGCGAGCGCAACATCGGCAAGGTCTTCAAGCATCGGTATACAACGACGAAACATCC 796
      |||...|||...|||...|||...|||...|||...|||...|||...|||...
Db      303 ----- 303
Qy      797 AGACACGGCTCTGGCTACCGGAAACCGTCAAGCATCGCCCTCTACCAAAAATA 856
      |||...|||...|||...|||...|||...|||...|||...|||...|||...
Db      304 -----GluThrileLeuLysProSerTyrrLeuSerProAla 315
      |||...|||...|||...|||...|||...|||...|||...|||...|||...
Qy      857 CAGACGGGACGATCGACTGCGAGCGTCCAAACCGCCGTGATCAGGACACCCCTGCA 916
      |||...|||...|||...|||...|||...|||...|||...|||...|||...
Db      316 LeuThrAsnMetLysAspAspGlyGlyValGln-----MetileAlaSerGlyPro--- 332
      |||...|||...|||...|||...|||...|||...|||...|||...|||...
Qy      917 AAAGGACTCGACCCGCGACCGCGCAACACCGCGATTGCGGGAGGCTGCAAGACGCGCTC 976
      |||...|||...|||...|||...|||...|||...|||...|||...|||...
Db      333 -----AspThrArgTrpGluileAsnLysLysLeuPheSerMetile 347
      |||...|||...|||...|||...|||...|||...|||...|||...|||...
Qy      977 AAACAGCCGAAAAAGCGTATCTGTTTCACCTATTTCGTCCTACAAAATCCGGC 1036
      |||...|||...|||...|||...|||...|||...|||...|||...|||...
Db      348 ThrSerAlaLysLysSerileTrpIleThrSerProTyrrPheileProAspGluAsp 367
      |||...|||...|||...|||...|||...|||...|||...|||...|||...
Qy      1037 ACAGACCACTGGCAAACTGGTCGAGGCGCATACAGCTTACCGTCTCTGACCACTCG 1096
      |||...|||...|||...|||...|||...|||...|||...|||...|||...
Db      368 LeuSerAlaLeuLysileAlaAlaLeuSerGlyileAspValArgileLeuValProAsn 387
      |||...|||...|||...|||...|||...|||...|||...|||...|||...
Qy      1097 CTACAGCGACCGAGCTTGGCGCGTCCATTCGCGCTACGTCAATACCGAAACCGCTG 1156
      |||...|||...|||...|||...|||...|||...|||...|||...|||...
Db      388 -----ArgProAspLysArgileValPheHisAlaSerArgSerTyrrPheProGluLeu 405
      |||...|||...|||...|||...|||...|||...|||...|||...|||...
Qy      1157 CTCAAAGCGGATCAAACTTACGAGCTGCAACCCCAACCATCGCTCCCGCCGACAAAA 1216
      |||...|||...|||...|||...|||...|||...|||...|||...|||...
Db      406 LeuGluAlaGlyValLysValTyrrGluTyrr----- 415
      |||...|||...|||...|||...|||...|||...|||...|||...|||...
Qy      1217 GACAAAGCGCTGACCGCAGCTCGTAAACAGCGCTGATGCCAAACCTTCATTGTGGAC 1276
      |||...|||...|||...|||...|||...|||...|||...|||...|||...
Db      416 ThrArgGlyPhe-----LeuHisSerLysileileileValAsp 428
      |||...|||...|||...|||...|||...|||...|||...|||...|||...
Qy      1277 GGCAAGCGCATCTTCATCGGCTCATCAACTCGACCCCGTTCGCGACCGCTCAATACC 1336
      |||...|||...|||...|||...|||...|||...|||...|||...|||...
Db      429 AsnGluileAlaSerileGlyThrSerAsnMetAspMetArgSerPheHisLeuAsnPhe 448
      |||...|||...|||...|||...|||...|||...|||...|||...|||...
Qy      1337 GAAATGGCGTCTCATCGAAAGCCCAAAATCGCAGAACATGGAGCGACCCCTCGCC 1396
      |||...|||...|||...|||...|||...|||...|||...|||...|||...
```

```
Db      449 GluValAsnAlaPhe-----LeuTyrrArgThrLysSer 459
Qy      1397 GATACCAACACCCGAAATAGCTACCGGTTCACCTTCGACAAACACACCGCTGCAATGG 1456
      |||...|||...|||...|||...|||...|||...|||...|||...|||...
Db      460 ValThrThrLeuValSerAspPheValTyrrAspLeuGluHisThrAsnGlnileArgPhe 479
      |||...|||...|||...|||...|||...|||...|||...|||...|||...
Qy      1457 CACGATCCCGCCACCCGAAACCTAC 1483
      |||...|||...|||...|||...|||...|||...|||...|||...|||...
Db      480 GluGlnPheArgAsnArgAlaTrpTyrr 488
      |||...|||...|||...|||...|||...|||...|||...|||...|||...

RESULT 28
C84125
cardiolipin synthetase BH3803 [imported] - Bacillus halodurans (strain C-125)
C:Species: Bacillus halodurans
C:Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 09-Jul-2004
C:Accession: C84125
R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Fujii, F.; Hir
Nucleic Acids Res. 28, 4317-4331, 2000
A:Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans an
A:Reference number: A83650; MUID:20512582; PMID:11058132
A:Accession: C84125
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-398 <STO>
A:Cross-references: UNIPROT:Q9K6C6; UNIPARC:UPI00000C42DC; GB:AP001520; GB:BA000004; N
A:Experimental source: strain C-125
C:Genetics:
A:Gene: BH3803
C:Superfamily: cardiolipin synthase

Alignment Scores:
Pred. No.: 1-1e-08 Length: 398
Score: 275.00 Matches: 108
Percent Similarity: 37.9% Conservative: 63
Best Local Similarity: 23.9% Mismatches: 172
Query Match: 9.6% Indels: 108
DB: 2 Gaps: 14

US-10-665-990A-13 (1-1561) x C84125 (1-398)
Qy      179 CGGACACCCCTCATACAAACGGGCTATCCGACATCTACCTGCTCGACGCCACCCACGAA 238
      |||...|||...|||...|||...|||...|||...|||...|||...|||...
Db      35 ArgAsnGlnProileArgHisGly-----ThrValArgLeuLeuProThrGlyAspGlu 52
      |||...|||...|||...|||...|||...|||...|||...|||...|||...
Qy      239 GCCTTTCGCCCGCGCGCTTATCGAATCTCCGACATCTCCGAGCTGTGTCAACCTCATGTACCTTGC 298
      |||...|||...|||...|||...|||...|||...|||...|||...|||...
Db      53 PheileGlySerLeuileHisAspilleLysAlaGluHisileHisileLeuPhe 72
      |||...|||...|||...|||...|||...|||...|||...|||...|||...
Qy      299 TACATTTGGCGCAACGACATTTCCGGCAGGCTGTGTCAACCTCATGTACCTTGC 358
      |||...|||...|||...|||...|||...|||...|||...|||...|||...
Db      73 TyriLePheArgAspHisileGlyLysileLysileLeuGlnHisLeuGlnLysAla 92
      |||...|||...|||...|||...|||...|||...|||...|||...|||...
Qy      359 GAACGGCGGTGCGCGTACGCTGCTGTGGAC-----GACAAACACACCGCGC 406
      |||...|||...|||...|||...|||...|||...|||...|||...|||...
Db      93 LysGluGlyValAlaValArgLeuLeuValAspArgPheGlyAlaAspValLysAsnArg 112
      |||...|||...|||...|||...|||...|||...|||...|||...|||...
Qy      407 GCGTTGGACGATCTCTGCTCGCCCTCGACAGCATCCCAATATCGAAGTGGCGCTTTC 466
      |||...|||...|||...|||...|||...|||...|||...|||...|||...
Db      113 SerileGlnSerLeuLysGlnAla---GlyAlaGlnPheGluTyrrAlaHisArgileSer 131
      |||...|||...|||...|||...|||...|||...|||...|||...|||...
Qy      467 AACCCCTTCCTCCTACGAAATGGCGGCACTCGGCTACCTGACCGGACTTCCCGCGCTC 526
      |||...|||...|||...|||...|||...|||...|||...|||...|||...
Db      132 PheProTyrr-----Trp-----PhePheSerLeu 139
      |||...|||...|||...|||...|||...|||...|||...|||...|||...
Qy      527 AACCGCGCATGCACAAATCCTTTACCGCGCAACACCGCGCCACCATCTACGCGGA 586
      |||...|||...|||...|||...|||...|||...|||...|||...|||...
Db      140 AsnArgArgAsnHisArgLysileThrValileAspGlyLysileGlyTyriLeGlyGly 159
      |||...|||...|||...|||...|||...|||...|||...|||...|||...
Qy      587 CGCAATATCGCGACGAAATCTTCAAAGTCGGTGAGGACACC-----GTTTTCGCC 637
      |||...|||...|||...|||...|||...|||...|||...|||...|||...
Db      160 TyrrAsnileGlyAspGluTyrr-----LeuGlyArgAspProLysLeuGlyPheTrpArg 177
      |||...|||...|||...|||...|||...|||...|||...|||...|||...
Qy      638 GACCTGGACATCTCGCCACCGGAGGCTCGTCGGCGAAGTATCGCAGCACTTCGACCGC 697
      |||...|||...|||...|||...|||...|||...|||...|||...|||...
```



```
Db 178 AspTyrHisLeuArgLeuThrGlyAspGlyValGlnAspLeuGlnAspGlnPheIleGln 197
Qy TACTGGCAAGCATTTCGCCGCACAAACCCACGCGCATCATCCGCGAGCGCAACATCGGC 757
Db 198 AspTrpGluArgGluSerArgLeuProValThrArg----- 209
Qy 758 AAGGCTCTTCAAGCACTCGGATACAAACGACGAAACATCCAGACACGCGCTCTCTGGGTAC 817
Db 209 ----- 209
Qy 818 CGCGAAACCGTGAACAGTCCGCCCTCTACCAAAAAATACAGCGGACGATCGACTGG 877
Db 209 ----- 209
Qy 878 CAGAGCGTCCAAACCGCGCTGATCAGCGACACACCCCTGCAAAAGGACTCGCGCGACGCG 937
Db 210 ----- 210
Qy 938 CGCAACCGCGATTGCGGAGGCTGCAAGCGG-----CTCAACAGAGCC 985
Db 226 IleIleProThrAsnGlySerPheLeuGluAspSerPheLeuGlnLeuValGluGlnAla 245
Qy 986 GAAAAAGCGTCTATCTGTTTCACCCCTATTTCGCTTACAAATCCGACGCGCACAGAGCA 1045
Db 246 GluGluThrIleThrIleGlyThrProTyrPheIleProGlyGluLeuHisAla 265
Qy 1046 CTGCAAAACTGGTCAGGCGGATAGACGTTACCGTCTGACCAACTCGCTACAGCG 1105
Db 266 LeuLeuAspAlaAlaArgGlyValAlaValArgLeuLeuValPro-----LysLys 283
Qy 1106 ACCGAGTTGCGCGCTGATCCGCGTACGTCAAATACCGAAACCGCTGCTCAAGCC 1165
Db 284 GlyAspHisProLeuValLysGluAlaAlaPheProTyrPheLysGluLeuGluGly 303
Qy 1166 GSCATCAAACTCTAGAGTGCAACCCACCATGCGTCCCGCCGACCAAAAGAGCAAGGC 1225
Db 304 GlyIleAsnIleTyrArgTyr----- 313
Qy 1226 CTGACCGCGAGCTCGGTAAACCGCTGATGCAAAACCTTCAATTGTGGAGCGCAACGC 1285
Db 314 Phe-----PheHisGlyLysAlaIleValIleAspLysLeu 326
Qy 1286 ATCTTCATCGGCTATTCACTGACCCCGTTCGCGACGCTCAATACGAAATGGCG 1345
Db 327 AlaAspValGlyThrAlaAsnPheAspLysArgSerPheArgLeuAsnTyrGluIleAsn 346
Qy 1346 GTCGTCATCGAAAGCCCAAAATCGCAACAGATGAGCGCACCTCGCGATACCACA 1405
Db 347 CysLeuLeuTyrAspLysGluMetIleGlnLeuValArgGluGluLeu----- 362
Qy 1406 CCCGAATACGCTACCGGTTTACCCTCGACAAACAAACCGCTGCAATGGACGATCCC 1465
Db 363 ---AspTyrAspPheSerIleSer-----GluArgLeuGlnMetGluAspLeu 377
Qy 1466 GCCACCCGAAACCTTACCCGACCAACCCGAA 1498
Db 378 AlaAsnArgSerPhePheHisArgThrLysGlu 388
```

RESULT 29

T43863

cardiolipin synthase [imported] - Clostridium perfringens

C:Species: Clostridium perfringens

C>Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 09-Jul-2004

C:Accession: T43863

R:Koyama, M.

submitted to the EMBL Data Library, August 1998

A:Description: Clostridium perfringens hem operon.

A:Reference number: Z22707

A:Accession: T43863

A:Status: preliminary; translated from GB/EMBL/DDBJ

A:Molecule type: DNA

A:Residues: 1-476 <KOY>

A:Cross-references: UNIPROT:Q9ZNC6; UNIPARC:UPI000016EODF; EMBL:AB017186; PIDN:BAA74786.
A:Experimental source: strain NCTC8237; vegetative cell
C:Genetics:
A:Gene: cls
C:Superfamily: cardiolipin synthase

Alignment Scores:
Pred. No.: 1.46e-08 Length: 476
Score: 273.00 Matches: 108
Percent Similarity: 42.1% Conservative: 79
Best Local Similarity: 24.3% Mismatches: 167
Query Match: 9.6% Indels: 90
DB: 2 Gaps: 20

US-10-665-990A-13 (1-1561) x T43863 (1-476)

```
Qy 107 CTGGAAGAACCGGCGGAAAGCCGTCATTTCATATCAATCTTCCAAACCT---GTCTCTCTGCAC 163
Db 66 LeuAspAspLysThrLysGlnLysTyrLeuAsnSerPheLysSerHisTyrLysLeuAsp 85
Qy 164 AACATC-----CTGCAATCCGGCACACCCCTCATCAACACGCGGCTA 205
Db 86 AsnIleSerLeuLysTyrLysAspLeuIleMetMetAsnPheAsnAspAsnSerThr 105
Qy 206 -----TCGACATCTACTGCTCGACGACCCCGACGAGCCCTTGCAGCCCGC 253
Db 106 TyrThrGlnArgAsnAspIleAspLeuTyrPheAspAlaAsnSerLeuPheGluMet 125
Qy 254 GCCGCCCTTATCGAATCTGCGGAACACACGCGTCGATTGCAATACTACATTGGCGCAAC 313
Db 126 IleAspGluIleAsnLysAlaGluLysPheIleHisMetGluPheTyrIlePheLysSer 145
Qy 314 GACATTTCCGCGACGCTGCTTTCAACCTCATGATCCTTGGCGGAGAACGCGGTGCGC 373
Db 146 AspGluIleGlyLysLysIleLeuGlnAlaLeuThrLysLysAlaLysGluGlyValGlu 165
Qy 374 GTACGCTGCTGTGGAGGAC-----AACACACGCGCGGTTGGAC---GATCTCTGTG 424
Db 166 ValLysLeuLeuValAspSerIleGlyAsnSerIleHisLysLysAspIleAspLysLeu 185
Qy 425 CTGCGCTTCGACAGCCATCCCAATATGAAAGTGCGCTGTTCACACCCCTTCGTCCTACGC 484
Db 186 LysAlaAlaGlyGly-----AspPheLysIlePhePheProGlyPheCysLys 201
Qy 485 AAATGGCGCGACCTCGGCTACTGACGACTTCCCGCTCAACCGCGGATGCACAC 544
Db 202 -----TyrIleAsn-----LeuArgIleAsnTyrArgAsnHisArg 213
Qy 545 AAATCTTTTACCGCGCACACCGCCACCATATCTCGGCGGAGCAAAATATCGGCGCAAC 604
Db 214 LysIleLeuIleIleAspSerLysValAlaPheLeuGlyGlyPheAsnIleGlyAspGlu 233
Qy 605 TACTTTCAAGTGGTAGGACACCGCTTTTCGCGGACCTGGACATCTCTCGCCACCGCGACG 664
Db 234 Tyr-----LeuGlyLysAspLys----- 239
Qy 665 GTCGTCGCGAAGTATCGCAGCACTTCGACGCTACTGGGCAAGCCATTCGCGCCACAC 724
Db 240 -----AsnIleGlyHis-----ArgAspThrHisThr 249
Qy 725 GCCAGCGCATCATCCGCGCGGCAACATC---GGCAAGGCTCTTCAAGCACTCGGATAC 781
Db 250 LysIleLysGlyLeuAlaIleAsnAspLeuGluGlyArgPheLeuLeuAspTrpSerTyr 269
Qy 782 AACGCAAAACATCCAGACACGCG---CTCTGCGCTACCGGAAACCGTGAACAGTCG 838
Db 270 AlaAsnGluSerAspLeuAspIleAspLysLysTyrPheIleAsnProHisSerThr 289
Qy 839 CCCCTCTACCAAAATAATACACGCGGACGATCGACTGGCAGAGCGTCCAAACCCGCTG 898
Db 290 AspLeuProLysLysIle-----IleGlyAlaGlnIle 300
Qy 899 ATCAGCGACACCCCTGCAAAAGGACTCGACGCGGACCGCGCAACACCGCGGATTCGCGG 958
```



```
Db 371 ArgAlaGlyIleLysIleAlaLeuTyArgProHisPhe----- 383
Qy 1220 AAAGCCTGACCGGAGCTCCGTAAACAGCTGCGATGCAAAACCTTCATTGTGTGACGGC 1279
Db 384 -----LeuHisAlaLysHisLeuSerValAspAsp 393
Qy 1280 AAACGCATCTTCATCGGCTCATTCACCTCGACCCCGTTCGGCAGCGCTCAATACCGAA 1339
Db 394 AspIleAlaLeuPheGlyIleAsnLeuAspIleArgSerPheAlaLeuAsnAlaGlu 413
Qy 1340 ATGGCGTGTGTCATCGAAAGCCCAAAATCGCAGACAGATGGAGCGCACCTCGCCGAT 1399
Db 414 IleGlyMetIleCysTyArgSerArgIleValSerArgLeuGluValGluSerAsp 433
Qy 1400 ACCACCCGAATAGCCTACCGGCTTACCTCGACAAACACACACCGCTGCAATGGCAC 1459
Db 434 -----TyrLeuAlaAsnAlaCysGlnLeuAsnLeuValGluTrpGln 447
Qy 1460 GAT---CCGCGCACCCGAAAAACC 1480
Db 448 GlnArgProAlaTrpArgArgSer 455

RESULT 31
T02345
hypothetical protein KIAA0324 - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 05-Mar-1999 #sequence_revision 05-Mar-1999 #text_change 09-Jul-2004
C:Accession: T02345
R:Ricke, D.O.; Bruce, D.; Mundt, M.; Doggett, N.; Munk, C.; Saunders, E.; Robinson, D.;
re, J.; White, S.; Ueng, S.; Tatam, O.; Campbell, C.; Fawcett, J.; Deaven, L.
submitted to the EMBL Data Library, March 1998
A:Description: Sequencing of human chromosome 16p13.3.
A:Reference number: Z14664
A:Accession: T02345
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1791 <RIC>
A:Cross-references: UNIPROT:060382; UNIPARC:UPI00000730AE; EMBL:AC004493; NID:g2996648;
C:Genetics:
A:Map position: 16
A:Introns: 1610/2; 1706/2
A>Note: KIAA0324

Alignment Scores:
Pred. No.: 2,97e-08 Length: 1791
Score: 268.50 Matches: 131
Percent Similarity: 37.6% Conservative: 51
Best Local Similarity: 27.1% Mismatches: 217
Query Match: 9.4% Indels: 85
DB: 2 Gaps: 18

US-10-665-990A-13 (1-1561) x T02345 (1-1791)
Qy 172 GCAATCCGGCACACCCCTCATACACACGGGCTATCCGACTACCTGCTCGACGACC 231
Db 689 SerSerProGluHisProProlysSerArgThrAlaArgGlySerArgSerPro 708
Qy 232 ---CCACGAAGCCCTTGGCGC-----CCGCGCCCGCTTATCGAATCTGCGGAACACAG 282
Db 709 GluProLysThrLysSerArgThrProProArgArg-----ArgSerArgSerSer 726
Qy 283 CTTCAATGTCATACTACTATTTGGCGCAACGACATTTCCGCGAGCTGCTGTTCAACCT 342
Db 727 ProGluLeuThrArgLysAlaArgLeuSerArgArgSerArgSerAlaSerSerPro 746
Qy 343 CATGTACCTTGGCGAGA---ACGCGCGTGGCGGTACGCTGCTGTTGGACGACAA 399
Db 747 GluThrArgSerArgThrProProArgHisArgArgSerProSerValSerSerProGlu 766
Qy 400 CACGCGCGGTGGACGATCTCTCTCGCTCGCCCTCGACAGCCATCCCAATATCGAATGGG 459
Db 767 ProAlaGluLysSerArgSerArgArgArgArgSerArgSerArgSerProArgThr--- 785
```

```
Qy 460 CCTGTTCAACCCCTTCCTCTACGCAATGCGCGCACTCGGCTACCTGACCGACTTCCC 519
Db 786 -----LysThrThrSerArgArgGlyArgSerPro 795
Qy 520 CCGCTCAACCGCGCATGACAAACAATCTTTACCGCCGACCAACCGCGCCACCATACT 579
Db 796 SerProLysProArgGlyLeuGlnArgSerArgSerArg---ArgGluLysThr 814
Qy 580 CGCGGACGCAATATCGCGACGAATACTTCAAGTGGTGAGACACCGCTTTTCGCCGA 639
Db 815 ArgThrThrArgArgAspArgSerGlySerGlnSerThrSerArgArgGln 834
Qy 640 CTTGGAACATCTCCGCCACCGCAGCTGTCGGCGAAGTATCGCAGCACTTCGACCGCTA 699
Db 835 ArgSerArgSerArgSerArgValThrArgArgArg 847
Qy 700 CTGGCGAAGCATCTCCGCCCAACCGCATATCCGACGCGCAACATCGGCAACATCGGCA 759
Db 848 -----GlyGlySerGlyThrHisSerArgSerProAlaArgGln 860
Qy 760 GGGTCTTCAACGACTCGGATACACGAGGAACATCCAGACCGGCTCCT-----CGG 813
Db 861 GluSerSerArgThrSerSerArgArgGlyArgSerArgThrProProThrSer 880
Qy 814 CTACGCGAACCCTCGAACAGTCCGCCCTCTACCA-----AAAAAT 855
Db 881 ArgLysArgSerArgSerArgThrSerProAlaProTrpLysArgSerArgSerArgAla 900
Qy 856 ACAGCGGGAGCATCGATCGATGGCAGCGTCCAAACCGCCTGATCAGCAGACCCCTGC 915
Db 901 SerProAlaThrHisArgSerArgSerArgThrProLeuIleSerArg----- 917
Qy 916 AAAGGACTCGACCGCGACCGCCGCAACCCGATTCGCGGAGGTGCAAGACCGCT 975
Db 918 ArgArgSerArgSerArgThrSerProValSerArgArgArgSerArgSerArgThrSer 937
Qy 976 CAACAGCCCAAAAGCGTCTATCTGTTTACCCTTATTCGTTCCCTCAAAAATCCGG 1035
Db 938 ValThrArgArgSerArgSerArgAlaSerProVal-----SerArg 952
Qy 1036 CACAGCGCACTGGCAAAACTGGTGCAGGACGGCATAGACTTACCTGCTGACCAATC 1095
Db 953 ArgArgSerArgSerArgThrProProValThrArgArgArgSerArgSerArgThrPro 972
Qy 1096 CTTACAGCGCGACAGCTTGGCGCGTCCATTCGGCTACCTCAAAATACCGAAACCGCT 1155
Db 973 ThrThr-----ArgArgSerArgSerArgThrProProValThrArg 987
Qy 1156 GCTCAAAGCCGCATCAAACCT-----CTACGAGCTGCAACCCAA 1194
Db 988 ArgArgSerArgSerArgThrProProValThrArgArgArgSerArgSerArgThrSer 1007
Qy 1195 CCATGCGCTCCCGCCCAAAAGAGCTGACCGGACAGCTCTTCATCGCTCATTAACCTCGACC 1254
Db 1008 ProlleThrArgArgArgSerArgSer-----ArgThrSerProVal 1021
Qy 1255 TGCCAAACCTTCATTGTGGACGCAACAGCATCTTCATCGCTCATTAACCTCGACC 1314
Db 1022 ThrArgArgSerArgSerArgThrSerProValThrArg----- 1035
Qy 1315 CGTTTCCGACGGCTCAATACCGAAATGGCGTCTCATCGAAAGCCCAAAATCGCAGA 1374
Db 1036 -----ArgArgSerArgSerArgThrSerProValThrArgArgArgSerArgSerArg 1053
Qy 1375 ACAGATGAGCGCACCTCTCGCGATACACACCGGAATACGCTTACCGGTACCTCGA 1434
Db 1054 ThrProProAlaIleArgArg-----SerArgSerArgThrProLeuLeuProArg 1071
Qy 1435 CAACACACCGCTGCAATGCGCATCCCGCACCCGCAAAACCTTACCCGACGAAC 1494
Db 1072 Lys-----ArgSerArgSerArgSerProLeuAlaIleArgArgSer 1086
Qy 1495 CGAGCCAAACTTTGGAAACGCATCGCGCAAAATCTCTATC-----CCTGCT 1542
```


A>Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1995
 R;Glaser, P.; Danchin, A.
 submitted to the EMBL Data Library, May 1995
 A;Description: Cloning and sequencing of the *Bacillus subtilis* chromosomal region from 3
 A;Reference number: S55414
 A;Accession: S55414
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 432-500 <GLA>
 A;Cross-references: UNIPARC:UPI000016B3F; EMBL:Z49782; NID:G853752; PIDN:CAA9861.1; PI
 R;Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berter
 C.; Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Ch
 A.; Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.
 Nature 390, 249-256, 1997
 A;Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gall
 tech, J.; Harwood, C.R.; Henauf, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.
 Koetter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois
 A;Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maue
 Y. M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle
 Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadale, Y.; Sato, T.; Scanlon,
 A;Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Seron
 akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama,
 T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yanane, K.; Yasumoto, K.; Yata, K.; Yoshida, K
 A;Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.
 A;Title: The complete genome sequence of the Gram-positive bacterium *Bacillus subtilis*.
 A;Reference number: A69580; MUID:98044033; PMID:9384377
 A;Accession: F70059
 A;Status: preliminary; nucleic acid sequence not shown; translation not shown
 A;Molecule type: DNA
 A;Residues: 1-500 <KUN>
 A;Cross-references: UNIPARC:UPI0000060B8CF; GB:Z99123; GB:AL009126; NID:G2636240; PIDN:CA
 A;Experimental source: strain 168
 C;Genetics:
 A;Gene: ywiE
 C;Superfamily: cardiolipin synthase

Alignment Scores:
 Pred. No.: 3.54e-08 Length: 500
 Score: 266.50 Matches: 107
 Percent Similarity: 35.8% Conservative: 64
 Best Local Similarity: 22.4% Mismatches: 176
 Query Match: 9.3% Indels: 131
 DB: 2 Gaps: 15

US-10-665-990A-13 (1-1561) x S60089 (1-500)

Qy	128	CGTCATTTCATACCTTCCAAACCTGCTCTCTGGACAAACATCTGCAAAATCGGCACACC	187
Db	94	ArgMetTyrAsnArgGluLys-----LeuArgLysLeuPheAspLysGluThr	110
Qy	188	CCTCATAACAAGGGCTATCCGAC-----	211
Db	111	ProGluValThrGlyLeuLysAspAsnGlnGluArgPhePheThrTyr-SerileArgAla	130
Qy	212	-----ATCTACCTGCTCGACGACCCCAAGCC	241
Db	131	AlaHisMetAsnIleAsnThrLysSerAsnIleLysValLeuLysAsnGlyGluThr	150
Qy	242	CTTGGCGCGCGCGCGCTTATCGAATCTGCGGAACACAGCTCGATTGCAATACTAC	301
Db	151	PheProAspIlePheIleAlaMetArgLysAlaGluSerTyrIleHisIleGluTyrTyr	170
Qy	302	ATTGGCGCAACGACATTTCGCGAGGCTGCTGTTCAACCTCATCTTACCTGCGCAGAA	361
Db	171	MetPheLysSerAspMetLeuGlyArgGlyMetMetAspIleMetGluLysAlaArg	190
Qy	362	CGCGCGCTGCGGTACGCTGCTGTGGACGACACACACGCGGGTGGACCATCTC	421
Db	191	GlnGlyValGluValArgPheLeuTyrAspAlaAlaGlySerMetLysLeuAlaArg	210
Qy	422	CTGCTCGCGCTCGACAGCCATCCCAATATCGAAGTGGCGCTGTTCAACCCCTCTCTCCTA	481
Db	211	AspIleMetArgMetLysGlnAlaGlyValAspIleValProPheSerProLeu-----	228

RESULT 34

VGBEX1

glycoprotein X precursor - equine herpesvirus 1 (strain Ab4p)

C:Species: equine herpesvirus 1

Qy	482	CGCAAAATGGCGCGCAGCTCGGCTACTACCGACGACTTCCCGCCGCTCAACCGCGCATGAC	541
Db	229	-----LysTyrGlyPheAsnGln-----LysLeuAsnPheArgAsnHis	242
Qy	542	AACAAATCCTTTACGCGCAACCGCGCACCATACTCGCGCGAGCGCAATATCGGCGAC	601
Db	243	ArgLysIleValIleIleAspGlyLysThrGlyPheValGlyGlyLeuAsnValGlyLys	262
Qy	602	GAATACTTCAAGTCGGTAGGACACCGCTTTTCCCGCAGCTTGGCAAGCATTCGCCGC	661
Db	263	GluTyr-----	264
Qy	662	AGCTGCTGCGCGAAGTATTCGACGAC-----TTCGACCGCTACTGGCAAGCATTCGCC	718
Db	265	-----IleSerArgAspProTyrIleGlyPheTyrArgAsp-----	276
Qy	719	CACAACGCCGCGCATCATCCGCGCGGCAACATCGGCAAGGTCTTCAAGCACTCGGA	778
Db	277	-----ThrHisLeuArgLeuGluGlyGluIleValGlnThrLeuHisAlaIlePhe	293
Qy	779	TACAACGAC-----GAAACATCCAGACACGCGCTCTCGCTACCGCGAA-----	823
Db	294	MetLeuAspTyrGluTyrValSerAsnGluValLeuIleAspGlnGluTyrAsnThr	313
Qy	824	-----ACCTGCGAAGTCGCCCTTACCAAAAATACAGCGGAGCATCGACTGG	877
Db	314	ProValProValGluGlyGlyIleTyrGlnIleValAlaThrGly-----	329
Qy	878	CAGAGCGTCCAAACCCGCGCTGATCAGCGACACCCCTGCAAAAGGACTCGACCGCGACGC	937
Db	330	-----ProAsp	331
Qy	938	CGCAACCGCGATTTCGCGGAGGCTGCAAGACGCGTCAAAACAGCCCGAAAAAGCGTC	997
Db	332	MetLysGluSerMetSerAspLeuTyrTyrGluMetIleSerSerAlaGlnLysSerile	351
Qy	998	TATCTGGTTTCACCTATTCCTCCATACAAAATCCGCGCACAGCGCACTGCAAAACTG	1057
Db	352	TrpIleAlaThrProTyrPheValProAsnGluSerIleArgThrAlaLeuLysAlaAla	371
Qy	1058	GTGCAGGACGCGATAGAGTTACCGTCTCGTACCAACTCGCTACAGCGCGACCGCTGCC	1117
Db	372	AlaThrLysGlyValGluValArgValMetValPro-----GluLysAsnAspSerPhe	389
Qy	1118	GCCGTCCATTCGCGCTACGTCAAAATACGAAACCGCTGCTCAAAAGCGCGCATCAACTC	1177
Db	390	LeuThrGlnTyrAlaSerArgSerTyrPheProGluLeuLeuGluGlyIleGluVal	409
Qy	1178	TACGAGCTGCAACCCACCATCGCTCCCGCCACAAAAGCAAGGCTGACCGCGCAGC	1237
Db	410	TyrSerTyrGln-----LysGlyPhe-----	416
Qy	1238	TCGCTAACCGCTGATGCCAAACCTTCATTGTGGACGCGCAACCGCATCTTTCATCGGC	1297
Db	417	-----MetHisGlnLysValMetIleIleAspGlyAspLeuAlaSerValGly	432
Qy	1298	TCATTCAACCTCGACCCCGCTTCCGCGACGCTCAATACGAAATGGGCGTCTGTCATC---	1354
Db	433	ThrAlaAsnMetAspMetArgSerPheGlnLeuAsnPheGluValAsnValPhePheThr	452
Qy	1355	-----GAAAGCCCC	1363
Db	453	AspAlaGluAlaIleArgThrLeuGluAlaHisPheGluGluAspMetGlnGluSerGlu	472
Qy	1364	AAATTCGCAAGACAG-----ATGAGGCGCGACCTTCGCGCATACACACCCGAA	1411
Db	473	LysLeuSerProValGlyPheTyrLysArgGlyValAlaAlaAspArgThrLysGlu	490

A>Note: host Equus caballus (domestic horse)
C>Date: 30-Sep-1992 #sequence_revision 30-Sep-1992 #text_change 09-Jul-2004
C/Accession: H36802
R:Telford, E.A.R.; Watson, M.S.; McBride, K.; Davison, A.J.
submitted to GenBank March 1992
A:Description: The DNA sequence of equine herpesvirus-1.
A:Reference number: A36805
A:Accession: H36802
A:Molecule type: DNA
A:Residues: 1-797 <TEL>
A:Cross-references: UNIPROT:P28968; UNIPARC:UPI0000138750; GB:M86664; NID:g330791; PIDN:
R:Telford, E.A.R.; Watson, M.S.; McBride, K.; Davison, A.J.
Virology 189, 304-316, 1992
A:Title: The DNA sequence of equine herpesvirus-1.
A:Reference number: A41831; MUID:92295566; PMID:1318606
A:Contents: annotation; possible protein-coding frames
A>Note: neither amino acid nor nucleotide sequence is given
C:Genetics:
A:Gene: 71
C:Superfamily: equine herpesvirus glycoprotein X; equine herpesvirus 1 glycoprotein homoc
C:Keywords: glycoprotein; transmembrane protein
F:1-22/Domain: signal sequence #status predicted <SIG>
F:23-797/Product: glycoprotein X #status predicted <MAT>
F:23-465/Region: serine/threonine-rich
F:489-797/Domain: equine herpesvirus 1 glycoprotein homology <EHG>
F:766-790/Domain: transmembrane #status predicted <TMN>
F:590/Binding site: carbohydrate (Asn) (covalent) #status predicted

Alignment Scores:
Pred. No.: 3.92e-08 Length: 797
Score: 266.00 Matches: 139
Percent Similarity: 36.2% Conservative: 52
Best Local Similarity: 26.3% Mismatches: 206
Query Match: 9.3% Indels: 131
DB: 1 Gaps: 20

US-10-665-990A-13 (1-1561) x VGBEX1 (1-797)

Qy	20	TCTGAAACTATATCCCGATGAAACACGACGACCTCATTTCCCTTTTATGCTCCTTCTC	79
Db	52	SerProThrThrSerProProThrThrSerSer-----	62
Qy	80	TGTTTCATGTTCTTCATGGTTGCCCTGACGAGAACGCGGAAAGCCGTCATTTCAT	139
Db	63	-----SerProProThrThrSerThrHisThrSerSerProSer---Ser	75
Qy	140	ACTTCCAAACCTGTCTCTCGGACAACTCTCGAAATCGGGACACCCCTCAT	194
Db	76	ThrSer-ThrGlnSerSerThrAlaAlaThrSerSerSerAlaProSerThrAlaSe	95
Qy	195	-----ACAAAGGGCTATCCGACATCTACCTGCTCGACGACCCCAAGCCCTTGCCGCC	250
Db	95	rSerThrThrSerThrSerThrSerThr-----SerThrGluThrThrThrThrProTh	114
Qy	251	CGGCGCCCTTATCGAATTCGCGAACACAGCCTCGATTTCGAATACTACATTGGCGC	310
Db	114	rAlaSer-----ThrThrThrProThrThrThrThrAlaAlaProThrThrAlaAlaTh	132
Qy	311	AACGACATTCGCGCAGCTGCTTCAACCTCATGTACCTTCCGCGAGAACCGGGGTG	370
Db	132	rThrThrAlaValThrThrAlaAlaSerThrSer-----	143
Qy	371	CGCGTACGCTGTGTGGACGACAAACACGCGGGTGGACGATCTCTGCTCGCC	430
Db	144	-----AlaGluThrThrThrThrThrAlaThrAlaThrAlaThrThrSerThrPr	157
Qy	431	CTCGACAGCATCCCAATATCGAAGTCGCGCTGTTCACCCCTTCGCTACGCAATGG	490
Db	157	oThrThrThrThrThrThrThrThrThrThrThrThrThrThrThrThrThrThrThr	177
Qy	491	CGGCACTCGGCTACCTGACGACTTCCCGGCTCAACCGCGCATGACACAAATCC	550
Db	177	rAlaSerThrThrThrThrThrThrThrThrThrThrThrThrThrThrThrThrThrAl	197

Qy	551	TTTACCGCGACACCGCGCACCATACTCGGCGGAGCGCAATATCGGCGCAACTACTTC	610
Db	197	aAlaThrThrThrAlaAlaThrThrThrAlaAlaThrThrThrAlaAlaThrThrAl	217
Qy	611	AAAGTCGGTGAGGACACCGTTCCTCGCGACCTCGACATCTCTCGCACCGCGAGGTCTC	670
Db	217	aAla-----ThrThrThrAlaAlaThrThrSerSe	227
Qy	671	GGCGAAGTATCGCAGGACTTCGACCGCTACTGGCAAGCCATTCCGCCCAACGCCACG	730
Db	227	rAlaThrThrAlaAlaThrThrThrAlaAla-----ThrThrThrAl	241
Qy	731	CGCATCATCGCAGCGCAACATCGGCAAGGTCTCAAGCACCTCGGATACAAACGCGAA	790
Db	241	aAlaThrThrThrAlaAlaThrThrThrAlaAlaThrThrThrAlaAlaThrThrGl	261
Qy	791	ACATCCAGACACGCGCTCTCGCTACCGGAAACCGTCGACACAGTCGCGCCCTCTACCA	850
Db	261	ySerProThrSerGlySerThrSerThrThrGlyAlaSerThrSerThrProSer----	279
Qy	851	AAAATACAGAGGAGCCATCGATGCGAGAGCGTCCAAACCCCTGTATCAGGACACC	910
Db	280	-----AlaSerThr-----AlaThrSerAlaThrPr	288
Qy	911	CCTGCAAAAGGACTCGACCGCGACCGCGCAACCGCGGATTGCGGGAGGCTCAAGAC	970
Db	288	oThrSerThrSerThrSerAlaAlaThr-----	298
Qy	971	CGCGTCAACAGCCCGAAAGCGTCTATCTGTTTTCACCCATTTCTGCTCCACAAA	1030
Db	299	-----ThrSerThrProThrProThrSerAlaAlaThrSerAlaGluSerThrGluAl	317
Qy	1031	TCGCGCACACGACGACTGCGCAAAACTGGTGAGGACGCGATAGACGTTACCGTCTGACC	1090
Db	317	aProThrSerThr-----	322
Qy	1091	AACTCGCTACAGCGCGACCGCTTTCGCGCTCCATTTCGCGTACGTCAAAATACGAAA	1150
Db	322	oThr-----ThrAspThrThrThrProSerGluAlaThrThrAlaThrThrSerPr	339
Qy	1151	CCGCTGCTCAAGCCCGCATCAAACTCTACGAGCTGCAACCCCAACCATG-----	1199
Db	339	oGluSerThrThrValSerAlaSerAlaThrThrSerAlaThrThrAlaPheThrGl	359
Qy	1200	-----CCGTCGCGCCCAAAAGAACAAAGGCTGACCGGACGCTCGTAACCGCGCTG	1252
Db	359	uSerHisThrSerProAspSerSerThr-----	368
Qy	1253	CATGCCAAACCTTCATTGTGGACGCGCAACGCGATCTTCATCG-----GCTCATTC	1303
Db	369	-----GlySerThrThrAlaGluProSerSerThrPheThrLeuThrProSe	385
Qy	1304	AACTCGACCCCGTTTCGCGACGCTCAATACCGAAATGGCGCTGTCATCGAAAGCCCC	1363
Db	385	rThrAlaThrPro-----SerThrAspGlnPheThrGlySerSerAlaSerTh	401
Qy	1364	AAAATCGCAGACAGATGGAGCGCACCTCGCGATACCAACCCGAATAGCCCTACCGC	1423
Db	401	rGluSerAspSerThrAspSer-----SerThrValProThrThrGlyThrGluSerIl	419
Qy	1424	GTTACCTCGCAAAACACACCGCGCTGC-----AATGCGACGATCCCGCC	1468
Db	419	eThrGluSerSerSerThrThrThrGluAlaSerThrAsnLeuGlySerSerThrTyrGlus	439
Qy	1469	ACCCGAAAAACCTACCCGACGAAACCGAGCGCAAACTTTTGGAAACGATCGCGCGAAA	1528
Db	439	rThrGluAlaLeuGluThrProAspGlyAsnThrThrSerGlyAsnThrThrProSerPr	459
Qy	1529	ATCCTATCCCTGCTGCCCATCG	1550
Db	459	oSer---ProArgThrProSer	465

RESULT 35
AE3539
cardiolipin synthetase (EC 2.7.8.-) [imported] - Brucella melitensis (strain 16M)
C:Species: Brucella melitensis
C>Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 09-Jul-2004
C:Accession: AE3539
R:DelVecchio, V.G.; Kapural, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Ios, T.; Ivanova,
.; Mazur, M.; Goltzman, E.; Selkov, E.; Elizer, P.H.; Hagius, S.; O'Callaghan, D.; Letess
Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002
A:Title: The genome sequence of the facultative intracellular pathogen Brucella melitensis
A:Reference number: AD3252; PMID:11756688
A:Accession: AE3539
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-492 <KUR>
A:Cross-references: UNIPROT:Q8YD8; UNIPARC:UPI00000583DD; GB:AE008918; PIDN:AAL53480.1;
A:Experimental source: strain 16M
C:Genetics:
A:Gene: BME10239
A:Map position: II
C:Superfamily: cardiolipin synthase
C:Keywords: transferase

Alignment Scores:
Pred. No.: 5,3e-08 Length: 492
Score: 263.50 Matches: 98
Percent Similarity: 39.2% Conservativeness: 69
Best Local Similarity: 23.0% Mismatches: 180
Query Match: 9.2% Indels: 79
Gaps: 13

US-10-665-990A-13 (1-1561) x AE3539 (1-492)

```
QY 128 GGTCAATTTCAATACATCTCAAAACCTCTCTCTGGACAACTCTGCAAAATCGGCACACC 187
Db 106 ArgGlnPheGlyAlaMetLys---lleLeuGlyAspAlaValSerLeuTyTrpAspPheThr 124
QY 188 COTCATAACAAGGGGTATTCGACATCTACCTGCTCGACGACCCCAAGAGCCCTTGGC 247
Db 125 SerGlyAsnThr-----lleGluMetLeuGluAlaGlyAspGluAlaTyrAla 140
QY 248 GCCCGCGCGCCCTTATCGAATCTCCGACACACGCTCGATTTGCAATACATACATTTGG 307
Db 141 AlaMetLeuGlyAlaIleGlyAlaGluArgSerIleValLeuGluThrTyrIlePhe 160
QY 308 CGCAACGACATTTCCGGCAGGCTGTCTTCAACCTCATGTACCTTCCCGCAGAACCGGC 367
Db 161 AspHisAspAlaIleGlyLysPheAlaAspAlaLeuGlyAspAlaValGlnArgGly 180
QY 368 GTGCGGTACGCTGCTGTGGACGACCAACACACGCGGGGTGGACGATCTCTGCTC 427
Db 181 ValGluValArgValLeuValAspAlaValGlyAlaArgTyrSerPheProSerIleVal 200
QY 428 GCCCTCGACAGCCATCCCAATATCAAGTGCCTTCAACCCCTTCTGCTCTACGCAA 487
Db 201 LysLeuLeuLysGluLysGlyValLysValAlaValPheAsnGlyAsnIleIleGly 220
QY 488 TGGCGCGCCTACCTGACCTGACCTTCCCGCCTCAACCGCGCATGCAACACAA 547
Db 221 LeuArg-----LeuProTyrAlaAsnLeuArgThrHisArgLys 233
QY 548 TCCTTTACGGCGACACACCGCGCCACCATCTCTCGGGGACGCAATTCGGGACCAATAC 607
Db 234 MetLeuIleValAspGlyLysThrAlaPheThrGlyGlyMetAsnIleArgAlaGlyPhe 253
QY 608 TTCAAA---GTGCGGTGAGGACACCGCTTTTCGCGACCTTGGACATCTCTCGCCACCGGACG 664
Db 254 ValArgAlaIleAlaGlyAspAlaValAlaPheAspThrHisPheLysLeuGluGlyPro 273
QY 665 GTCTGCGCGCAAGTATCGCAGCATTCGACCGCTACTGGCAAGCCATTCGCCCCCAAC 724
Db 274 AlaIleAlaAspLeuPheHisIleAlaSerGluAspTrpArg----- 287
```

```
QY 725 GCCAGCGCATCATCCGACGGCGCAACATCGGCAAGGGTCTTCAAGCACTCGGATACAA 784
Db 288 -----PheAlaThrGlyGluLeuLeuThrGly----- 296
QY 785 GACGAAACATCCAGACAGCGCTCTCGCTACCGCAAAACCGTCGAACAGTCGCGCCCTC 844
Db 297 -----GluAlaTrpSerIleAlaProPro 304
QY 845 TACCAAAAATACAGACGGGACGATCGATCGTGGCAGAGCGTCCAAACCCGCTGATAGC 904
Db 305 GluAsnProProGlyThrGlyThrLeu-----ValArgValValGly 318
QY 905 GACACCCCTGCAAAAGGACTCGACCGGACCGCGCAAAACCGCGGATTCGCGGAGGTG 964
Db 319 SerGlyProAspLysAsnLeuGluThrAsnHisArg-----MetMet 332
QY 965 CAAGACGCGCTCAAAACAGCCGCAAAAGCGCTCTCTGTTTTCACCTATTTCTGCTCCT 1024
Db 333 MetGlyAlaPheSerIleAlaGlnHisIleLeuIleMetThrProTyrLeuLeuPro 352
QY 1025 ACAAATCCGACAGACGCTGCGCAAACTGTGTGAGGACGATAGAGTTCACCTC 1084
Db 353 AspArgGluLeuIleSerAlaLeuValThrAlaAlaArgGlyValSerValAspIle 372
QY 1085 CTG-----ACCACTCGCTACAGCGACCGACGCTGCGCGCTCATTCGCGCTAC 1135
Db 373 ValValProGlyValAsnAsnLeuLysLeuValAsp-----ArgAlaMetArgAlaGlnPhe 391
QY 1136 GTCAATATCCGAAACCGCTGCTCAAAACCGCGCATCAAACTCTACGAGCTGCAACCCCA 1195
Db 392 -----AspGlnLeuLeuArgAspGlyCysArgIleTrpArg----- 403
QY 1196 CATCGCTGCCCGCCCAAAACAGAAAGCTGACCGGACGCTCGTAACAGACCTGCAT 1255
Db 404 -----AlaGlyGlyAlaPheAsn-----His 410
QY 1256 GCCAAACCTTCATTGTGGAGCGCAACCGCATCTTCATCGGCTCATCAACCTCGACCCC 1315
Db 411 SerLysLeuMetThrIleAspGlyAlaTrpSerTyrValGlySerSerAsnIleAspPro 430
QY 1316 GCTTCGCGCAGCTCAATACCGAAATGGCGCTGCTCATCGAAAGCCCCCAAAATTCGAGAA 1375
Db 431 ArgSerLeuArgLeuAsnPheGluValAspLeuGluIleLeuAspArgAspValAlaArg 450
QY 1376 CAGATGAGCGCACCTC 1393
Db 451 GlnValGluGluArgIle 456
RESULT 36
E97307
probable cardiolipin synthase (phospholipase D family) [imported] - Clostridium acetobu
C:Species: Clostridium acetobutylicum
C>Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 09-Jul-2004
C:Accession: E97307
R:Noiling, J.; Breton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee,
.; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.
J. Bacteriol. 183, 4823-4838, 2001
A:Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Cl
A:Reference number: A96900; MUID:21359325; PMID:21359325
A:Accession: E97307
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-510 <KUR>
A:Cross-references: UNIPROT:Q97B04; UNIPARC:UPI000000D756C; GB:AE001437; PIDN:AAK81248.1;
A:Experimental source: Clostridium acetobutylicum ATCC824
C:Genetics:
A:Gene: CAC3316
C:Superfamily: cardiolipin synthase
Alignment Scores:
Pred. No.: 6.08e-08 Length: 510
Score: 262.50 Matches: 119
Percent Similarity: 36.5% Conservativeness: 82
```


DB:	2	Gaps:	10
US-10-665-990A-13	(1-1561) x D83103 (1-359)		
Qy	263	ATCGAATCTCCGACACACAGCCTCGATTTCGAATCTACATTTTGGCCACAGCATTTCC	322
Db	5	IleGluAlaGlnAArgSerIleGluGluLeuTyrLeuValGluAspGlyHisCys	24
Qy	323	GGCAGGTGCTGTTTCAACCTCATGACTCTTCCGCGCAGAACCGCGCTGCGCTACGCCTG	382
Db	25	AlaGluLeuPheLeuValAlaLeuLeuAspAlaArgArgGlyValAlaValArgCys	44
Qy	383	CTGTTGGACGACAAACACACGCGGCTGTGACGATCTCTGCTCGCCCTCGACAGCCAT	442
Db	45	LeuPheAspGlyPheGlyCysLeuGlyLeuGlySerAlaTrpIleGlnAArgLeuAArgGlu	64
Qy	443	CCCAATATCGAAGTCGCGCTGTTCAACCCCTTCTCTACGCAAAATGGCGGCACTCGGC	502
Db	65	AlaGlyGlyGluLeuAArgLeuTyrAsnProLeu-----ArgTrpLysLeuThrGly	81
Qy	503	-----TACCTGACCGACTTCCCGCGCTCAACCGCGCATGACCAACAATCCTTT	553
Db	82	GlyAsnLeuTyr-----ArgAspHisAArgLysLeuLeu	92
Qy	554	ACCGCCGACAAACCGCGCACCATACTCGGGGAGCAATATCGGCGACGAATACTTCAAA	613
Db	93	LeuValAspGlyAArgLeuGlyTyrValGlyGlyAlaGlyIleThrAspGluPheTrpGlu	112
Qy	614	---GTGGTGAGGACACGGTTTTCCGCGACCTGGACATCTCGCCACCGGACGGTCTGTC	670
Db	113	ProValSerAspValSerAlaTrpAArgGluValMetValGluMetAspGlyProValVal	132
Qy	671	GGCGAAGTATCGCACGACTCGACCGTCTACTGGGCAAGCCATTCCGCCCAACAGCCG	727
Db	133	AlaAspTrpAlaAlaLeuPheGluArgGlnTrpLeuAlaCysLeuGluGluLysAlaTrp	152
Qy	728	-----ACGGCATCATCTCGCACGCGGCAACATCGGCAAGGGT	763
Db	153	LysProArgGluGlyMetThrLeuThrAArgLeuProGlnProGlyAlaAlaAArgGly	172
Qy	764	CTTCAAGCATCTGGATACACAGCAGAAACATCTCAGACACGGCGTCTCGCGTACCGCAA	823
Db	173	LeuGlyArgValAlaTyrAlaAsp-----	180
Qy	824	ACCTCGAACAGTCCGCCCTCTACCAAAATACAGACGGAGCATCGATGCGACAGC	883
Db	181	-----Ala	181
Qy	884	GTCCAAACCCCGCTGATCAGCAGACACCCCTGCAAAAGGACTTCGACGCGACCGCGCAA	943
Db	182	ArgGlnHisAArgAspIleLeuGlnSerLeuValArgAlaLeuAenGlySerAArgArg	201
Qy	944	CCGCGGATTCGGGAGGCTGCAAGACGCGTCAACAGCCCGCAAAAGCGTCTATCTG	1003
Db	202	-----IleTrpLeu	204
Qy	1004	GTITTCACCTATTTTCTGCTTACAAAATCCGGCACAGACGCACTGGCAAAACTGGTCAG	1063
Db	205	AlaThrProTyrPheLeuLeuProTrpTrpLysValArgArgAlaLeuAArgLysAlaAlaGln	224
Qy	1064	GACGGCATAGACGTTCACGCTCTGACCAACTCGGTACAGGCGACCGACGTTGCCGCGCTC	1123
Db	225	ArgGlyValGluValArgLeuLeuLeuAlaGly---ArgIleThrAspHisAlaProVal	243
Qy	1124	CATTCCGGCTACGTCAAATACCGAAAACCGGTGCTCAAGCCGCGCATCAAACTTACGAG	1183
Db	244	ArgTyrAlaGlyGlnAArgTyrTyrProArgLeuLeuAArgAlaGlyValArgIleHisGlu	263
Qy	1184	CTGCAACCCCAACCATGCGCTCCCGCCACAAAGACAAAGCCCTGACCGGACGCTCCGTA	1243
Db	264	TyrGlnProArgPhe-----	268
Qy	1244	ACCAGCTGCATGCGCAAAACCTTCATTGTGAGCGGCAAAACGATCTTTCATCGGCTCATTC	1303

Db 269 -----LeuHisLeuLysMetValMetValAsp---AspTrpValSerValGlySerCys 285
Qy 1304 AACCTGACCCCGTTCGCGACCGCTCAATACCGAAATGGCGTGCATCGAAAGCCCC 1363
Db 286 AsnPheAspHisTrpAsnLeuArgPheAsnLeuAspAlaAsnLeuGluAlaLeuAspPro 305
Qy 1364 AAAATCCGACAAACAGATGGAGCGCACCTC 1393
Db 306 AspPheThrAsnGluAlaAlaSerLeu 315
RESULT 38
AH2333
cardiolipin synthase [imported] - Nostoc sp. (strain PCC 7120)
C:Species: Nostoc sp. PCC 7120
A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Jul-2004
C:Accession: AH2333
R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchih
Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yasuda, M.; Tabata, S.; S
DNA Res. 8, 205-213, 2001
A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ana
A:Reference number: AB1807; MUID:21595285; PMID:11759840
A:Accession: AH2333
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-480 <KUR>
A:Cross-references: UNIPROT:Q8VPHI; UNIPARC:UPI00000CEA63; GB:BA000019; PIDN:BAB75922.1
A:Experimental source: strain PCC 7120
C:Genetics:
A:Gene: alr4223
C:Superfamily: cardiolipin synthase
Alignment Scores:
Pred. No.: 1.46e-07 Length: 480
Score: 256.00 Matches: 120
Percent Similarity: 38.3% Conservative: 69
Best Local Similarity: 24.3% Mismatches: 181
Query Match: 9.0% Indels: 124
DB: 2 Gaps: 18
US-10-665-990A-13 (1-1561) x AH2333 (1-480)
Qy 89 TCTTCATGTTGCCCGCCCACTGGAGAACGGACGGAAAGCGCTCATTTCAATACTTCCAAA 148
Db 98 SerAlaLysLeuGluProLeuGln----- 105
Qy 149 CTTGCTCTCTGGCAACATCATCTGCAAAATCGGGCACACCCCTCATACACACGGGCTATCC 208
Db 106 ---LeuLeuAlaGluAlaPheIleGlyLeProPheThrSerGlyAsnAsnAla----- 122
Qy 209 GACATCTACTGCTGCGACGACCCCGACGAGCCCTTCGCGCGCGCGCCCTTATCGAA 268
Db 123 -----LysLeuLeuIleAsnGlyGlnThrTyrAlaAlaMetLeuSerAlaIleAla 140
Qy 269 TCTGCCGAACACAGCCCTCGATTGTCGAATACTACATTTGGCGCAACGACATTTCCGGCAGG 328
Db 141 SerAlaAsnSerTyrIleLeuLeuGlnSerTyrIleValValAspAspLysAlaGlyAsn 160
Qy 329 CTGCTGTTCAACCTCATGTACCTTGGCGCAGAACGGCGGTGGCGGTACGCGCTGCTGTG 388
Db 161 GluPheLysAspAlaLeuIleAlaLysAlaLysGlnGlyIleArgValTyrLeuIleTyr 180
Qy 389 GACGACAAACACAGCGCGGGTGGAGCATCTCTGCTCGCCCTCGACGCCATCCCAAT 448
Db 181 AspGluIleGlySerAsnLysIleSerArgLeuTyrValLysSerLeuGlnLysTyrAsp 200
Qy 449 ATCGAAGTGGCGCTGTTTCAACCCCTTCGTCTACGCAATGGCGCGCACTCGGCTACCTG 508
Db 201 IleGlnValSerAlaPheHisThrArgGlyArgGlyAsnArgPhe----- 216
Qy 509 ACCGACTTCCCGCCTCAACCGCGCGCATGCAACAAATCTCTTACCGCGCGCAACCGC 568
Db 509 -----LeuHisLeuLysMetValMetValAsp---AspTrpValSerValGlySerCys 285


```
QY 576 TACTCGCGGACGCAATATCGCGGACGGAATACTTCAAAGTCGGTGAGGACACCGTTT-TC 634
Db 228 rProthrArgThrProProArgGluSerLeu : : : : : : : : : : : : : : : :
QY 635 GCGGACCTGGACATCTCGCCACCGCGAGCGTGGTGGGAGAGTATCGCAGCACTTCGAC 694
Db 245 rPro-----ThrArgMetProProArgAlaSerProThrArgProProArgAlaSe 263
QY 695 CGCTACTGGG-----CAAGCATTCCGCGCCACACGCGCAGCATCATCCGCGAGC 745
Db 263 rProThrGlySerProProArgAlaSerProMetThrProProArgAlaSerProArgTh 283
QY 746 GGCACATCGGCAAGGCTTCAACGACCTCGATACAGGACGCAACATCCAGACACGG 805
Db 283 rProProArgAla-----SerProThrThrThr-----ProSerAr 295
QY 806 CTCTCGCGTACCGGCAACCGTGAACAGTCGCGCCCTCTACCAAAAAATACAGA----- 860
Db 295 gAlaSerLeuThrArgThrProSerTrpAlaSerProThrThrThrProSerArgAlaSe 315
QY 861 -----CGGAGCGCATCGAC 874
Db 315 rLeuMetLysMetGluSerThrValSerIleThrArgThrProProArgAlaSerProTh 335
QY 875 TGGCAGAGGTCCAAACCGCGCTGATCAGCGACACCCCTGCAAAAGGACTCGACCGCAC 934
Db 335 rGlyThrProSerArgAlaSerProThrGlyThrProSerArgAlaSerLeuThrGlySe 355
QY 935 CGCCGCAAAACCGCATTCGCGGAGGCTGCAAGACGGCTCAACACGCGCCGCAAAAAAGC 994
Db 355 rProSerArgAlaSerLeuThrGly-----CGGAGCGCATCGAC 874
QY 995 GTCTATCTGTTTCAACCTATTTCGTCCTCAAAAATCCGCAACAGACGCACTGGCAAAA 1054
Db 368 aSerLeu-----IleGlyThrProSerArgAla----- 377
QY 1055 CTGGTGCAGGACGGATAGACGTTACCGTCTGACCAACTCGCTACAGCGCACCGACGTT 1114
Db 378 -----SerLeuIleGlyThrProSerArgAlaSerLeuThrGlyTh 391
QY 1115 GCGCGCG-----TCCATTTCGCGTACGTCAAAT-----ACCGAAAA 1150
Db 391 rProProArgAlaSerLeuThrGlyThrSerSerThrAlaSerLeuThrArgThrProSe 411
QY 1151 CGCTGCTCAAAAGCGCGCATCAAACTCTACGAGCTGCAACCAACCATCGCTCCCGCC 1210
Db 411 rArgAlaSerLeuThrArgThrGlnSerSerSerSerLeuThrArgThrProSerMetAl 431
QY 1211 ACAAAAGACAAGGCTGACCGGACGCTCCGTAACACGCTGCATGCCAAACCTTCATT 1270
Db 431 aSerLeuThrArgThrProProArgAlaSerLeuThrArgThrProProArgAlaSerLe 451
QY 1271 GTGGAGCGGAAACGCATCTTCATCGCTCATTCACACCTCGACCCCGTTCGCGACGGT 1330
Db 451 u-----ThrArgThrProProArgAlaSerLeuThrArgThrPro---ProArgAlaSe 468
QY 1331 AATA-----CCGAAATGGGCTCGTCATCGAAAGCCCAAAATCGCAGAACAGATG 1381
Db 468 rLeuThrArgThrProSerMetValSerLeuLysArgSerProSerArgAlaSerLeuTh 488
QY 1382 GAGCGCACCTCGCGCATACACACCGGAATACGCTACC-----GGTTACCTCGAC 1435
Db 488 rArgThrProSerArgAlaSerLeuThrMetThrProSerArgAlaSerLeuThrArgTh 508
QY 1436 AAACACAAACCGCTCGCAATGGCAGCATCCCGCCACCCGCAAAAAACCTACCGCAACGCC 1495
Db 508 rProSerThrAlaSerLeuThrGlyThrProProThrAlaSerLeuThrArgThrProPr 528
QY 1496 GAAGCCAAACTTTGGAAACGCATTCGCGCAAAAAATCCTATCCCTGCTGCCCATCG 1550
Db 528 othrAlaSerLeuThrArgSerProProThrAlaSerLeuThrArgThrProSer 546
```

RESULT 40

```
S21054
DNA-directed RNA polymerase (EC 2.7.7.6) II largest chain - human
N:Alternate names: DNA-directed RNA polymerase B largest chain
C:Species: Homo sapiens (man)
C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
C:Accession: S21054; S18987
R:Wintzerith, M.; Acker, J.; Vicaire, S.; Vigneron, M.; Kedinger, C.
Nucleic Acids Res. 20, 910, 1992
A:Title: Complete sequence of the human RNA polymerase II largest subunit.
A:Reference number: S21054; MUID:92178992; PMID:1542581
A:Accession: S21054
A:Molecule type: mRNA
A:Residues: 1-1970 <WIN>
A:Cross-references: UNIPROT:P24928; UNIPARC:UPI000000037D; EMBL:X63564; NID:g36123; PIDN
C:Genetics:
A:Gene: GDB:POLR2A; POLRA
A:Cross-references: GDB:120306; OMIM:180660
A:Map position: 17p13.1-17p13.1
C:Superfamily: human DNA-directed RNA polymerase II largest chain
F:Keywords: DNA binding; nucleotidyltransferase; tandem repeat; transcription; zinc finger
F:71-87/Region: zinc finger CCH motif

Alignment Scores:
Pred. No.: 3,4e-07 Length: 1970
Score: 250.50 Matches: 123
Percent Similarity: 38.2% Conservative: 69
Best Local Similarity: 24.5% Mismatches: 196
Query Match: 8.8% Indels: 115
DB: 1 Gaps: 24

US-10-665-990A-13 (1-1561) x S21054 (1-1970)
QY 108 TGAAGAAACGGAAGCCGTCATTTCATATCTTCCAAACCTGCTCCTCTGGACAACA 167
Db 1520 TrpAsnGlnGly-----AlaThrProAlaTyThrGlyAlaTrpSerPro 1533
QY 168 TCCTGCAAAATCGGCACACCCCTCATACAACGGGTATCCGACATCTACCTGC----- 221
Db 1534 SerValGlySerGlyMetThrProGlyAlaAlaGlyPheSerProSerAlaAlaSerAsp 1553
QY 222 -----TCGACGACCCCAACGAGCCCTTG 245
Db 1554 AlaSerGlyPheSerProGlyTyThrSerProAlaTrpSerProThrProGlySerProGly 1573
QY 246 CGCGCCGCGCGCCCTTATCGAATCTGCGAACAACAGCCTCGATTGCAATACTACTATT 305
Db 1574 SerProGlyProSerSerProTyTrpIleProSer----- 1584
QY 306 GCGCAACGACATTTCCGCGAGGCTGCTTCAACCTCATGTACCTTGC CGCAGAACGCG 365
Db 1585 -----ProGlyGlyAlaMetSerProSerTyThrSer---ProThrSerPro 1598
QY 366 GCGTGGCGTACGCTGCTGTTGGACGACAACAACGCGCGGGTTGGACGATCTCTCTGC 425
Db 1599 Ala-----TyrGluProArgSerProGlyGlyTyThrProGln--- 1611
QY 426 TCGCCTCGACAGCATCCGAATATCGAAGTGGCGCTGTTCAACCCCTTCGTCCTACGCA 485
Db 1612 SerProSerTyThrSer---ProThrSerProSerTyThrSerProSerTyThrSer 1630
QY 486 -----AATGGCGGCGACTCGCTACCTGACCGACT---TCGCCGCGCTCAACCGCGCA 536
Db 1631 ProThrSerProAsnTyThrSerProThrSerProSerTyThrSerProThrSerProSerTy 1650
QY 537 TGCACAAATAATTCCTTTTACCGCGCAACACCGCGCCACCATACTCGCGGAGCAATATCG 596
Db 1651 SerProThrSerProSerTyThrSerProThrSerProSerTyThrSerProThrSerProSer 1670
QY 597 GCGCAATACTTCAAGTCGGTAGGACACCGTTTTCGCCGACCTCGACATCTCTCGCA 656
Db 1671 TyrSerProThrSerProSerTyThr---SerProThrSerProSerTyThrSerProThrSer 1689
QY 657 CCGGCGAGCGTGC-----TCGGCGAAGTATCGACGACTTCGACCGCTACTCGGCAAGCC 710
```

```
Db 1690 ProSerTyrSerProThrSerProSerTyrSerProThrSerProThrSerProThr 1709
Qy 711 ATTCGCGCCACAAACCGCCGCGCATCATCCGCGGCAACATCGGCAAGGCTTCAAG 770
Db 1710 SerProSerTyrSerProThrSerProSerTyrSerProThrSerPro-----Ser 1726
Qy 771 CACTCGGATACAGACGAAACATCCAGACAGCGCTCTCTGCGTACCGCGCAAAACCGTCG 830
Db 1727 TyrSerProThrSerProSerTyrSerProThrSerProThrSerProThrSerPro 1746
Qy 831 AACAGTCGCGCTCTACCAAAAATACAGACGGGACGATCGACTGGCAGAGCGTCCAAA 890
Db 1747 AsnTyrSerProThrSerProAsnTyrThrProThrSerProSerTyrSerProThrSer 1766
Qy 891 CCCGCTGATCAGCGACACCCCTGCAAAAGGACTCGACCGCGCGCCGCGCAACCGCCGA 950
Db 1767 Pro-----SerTyrSerProThrSerProAsnTyrThrProThrSerProAsn----- 1782
Qy 951 TTGCGGGAGGCTGCAAGACGCGCTCAAAACAGCCGCAAAAGCGTCTATCTGGTTTCAC 1010
Db 1783 -----TyrSerProThrSerProSer-----TyrSer 1791
Qy 1011 CCTATTTCGCTCCTACAAAATCCGGCACAGACGCACTGGCAAAACTGGTGCAGGACGGCA 1070
Db 1792 ProThrSerProSerTyrSerProThrSerProThrSerProThrSerProThrSer 1805
Qy 1071 TAGAGTTACCGTCTCTGACCAACTCGCTAC-----AGCGACCGCGTTCGCGCGCTCC 1124
Db 1806 -----ProSerPro-----ArgTyrThrProGlnSerProThrTyrThrProSer 1821
Qy 1125 ATTCGCGCTAGGTCAATACCGAAACCGCTGCTCAAGCGCGCATCAAACTCTACGAGC 1184
Db 1822 SerProSerTyrSer-----ProSerSerProSerTyrSer 1833
Qy 1185 TGCACCCCAACGATCCGCTCCCGCCCAAAAGACAAAGGCGCTGACCGGACAGTCCGTA 1244
Db 1834 -----ProThrSerProLysTyrThr-----ProThrSerPro--- 1844
Qy 1245 CCAGCTGCATGCCAAACCTTCATTGTGGACGGCAACGCACTTCATCGGCTCATTC 1304
Db 1845 -----SerTyrSer 1847
Qy 1305 ACCTGACCCCGTTCCGACGCGCTCAATACCGAAATGGCG--TCGTATCGAAAGCC 1361
Db 1848 ProSerSerProGluTyrThrProThrSerProLysTyrSerProThrSerProLysTyr 1867
Qy 1362 CCAAAATCGAGACAGATGGAGGCACTTCGCGGATACCACACCGGAATACGCTACC 1421
Db 1868 SerProThrSerProLysTyrSerProThrSerProThrTyrSerProThrProLys 1887
Qy 1422 GCGTTACCTCGACAAACACACCGCTGCAATGGCAGATCCCGCCACCCGAAACCT 1481
Db 1888 TyrSerProThrSerProThrTyrSerProThrSerProValTyrThrProThrSerPro 1907
Qy 1482 ACCGGAACGAAACCGCAACCAACTTTGGAAACGATCGCGCAAAATCCTATCCCTGC 1541
Db 1908 LysTyrSerProThrSerProThrTyrSerProThrSerProThrSerProThrSer 1926
Qy 1542 TGCCCATCG 1550
Db 1927 SerProThr 1929
```

Search completed: May 2, 2006, 05:26:35
Job time : 143 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: May 2, 2006, 06:09:35 ; Search time 41 Seconds
(without alignments)
1232.044 Million cell updates/sec

Title: US-10-665-990A-14
Perfect score: 2720
Sequence: 1 MHTDPKIQAMPSETISPMKT.....KLMKRIAATLSLLPIEGLL 525
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues
Total number of hits satisfying chosen parameters: 283416
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 80.*
1: Pirl.*
2: Pirl2.*
3: Pirl3.*
4: Pirl4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2643	97.2	525	B81859	phospholipase D-fam
2	2573	94.6	508	B81083	cardiolipin syntha
3	1102.5	40.5	493	C64847	yndC protein - Esc
4	1100.5	40.5	493	H90806	probable synthase
5	1093.5	40.2	493	D85666	conserved hypothet
6	1089.5	40.1	495	AE0636	conserved hypothet
7	807	29.7	518	G97556	hypothetical 55.9K
8	807	29.7	518	AB2777	phospholipase D fa
9	792	29.1	529	F82983	conserved hypothet
10	632	23.2	502	E71963	probable cardiolip
11	627	23.1	502	F64543	conserved hypothet
12	439	16.1	652	B82724	cardiolipin syntha
13	314	11.5	482	G70063	cardiolipin syntha
14	308	11.3	494	F90001	hypothetical prote
15	300	11.0	504	A11433	cardiolipin syntha
16	294.5	10.8	504	A11075	cardiolipin syntha
17	292	10.7	482	A11762	cardiolipin syntha
18	287	10.6	482	AG1387	cardiolipin syntha
19	284.5	10.5	505	G89906	hypothetical prote
20	282.5	10.4	481	B86744	cardiolipin syntha
21	276	10.1	503	B84007	cardiolipin synthe
22	275	10.1	398	C84125	cardiolipin synthe
23	273	10.0	476	T43863	cardiolipin syntha
24	270.5	9.9	467	H82711	cardiolipin syntha
25	268	9.9	490	B82971	cardiolipin syntha
26	266.5	9.8	500	S60089	cardiolipin syntha
27	263.5	9.7	492	AE3539	cardiolipin synthe
28	261	9.6	359	D83103	probable phospholi
29	260	9.6	510	E97307	probable cardiolip

ALIGNMENTS

RESULT 1

B81859
Phospholipase D-family protein NMA1646 [imported] - Neisseria meningitidis (strain Z2491)
C;Species: Neisseria meningitidis
C;Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 09-Jul-2004
C;Accession: B81859
R;Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; More
; Holroyd, S.; Jagels, K.; Leather, S.; Moule, S.; Mungall, K.; Quail, M.A.; Rajandrea
Nature 404, 502-506, 2000
A;Title: Complete DNA sequence of a serogroup A strain of Neisseria meningitidis Z2491.
A;Reference number: A81775; MUID:20222556; PMID:10761919
A;Accession: B81859
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-525 <PAR>
A;Cross-references: UNIPROT:Q9JTT1; UNIPARC:UPI000005033F; GB:AL162756; GB:AL157959; NI
A;Experimental source: serogroup A, strain Z2491
C;Genetics:
A;Gene: NMA1646
C;Superfamily: cardiolipin synthase

Query Match	97.2%;	Score	2643;	DB 2;	Length	525;			
Best Local Similarity	97.3%;	Pred. No.	5.2e-182;						
Matches	511;	Conservative	5;	Mismatches	9;	Indels	0;	Gaps	0;
Qy	1	MHTDPKIQAMPSETISPMKTRSLISLLCLLCCSSWLPPLEERTESRHNTPSKVLLDN	60						
Db	1	MEANPKTQAMPSETISLMKTRSLISLLCLLCCSSWLPPLEERTESRHNTPSKVRLDN	60						
Qy	61	ILQIRHTPHNGLSDIYLLDDPHEAALAAALIESAHSLSLDQYIWRNDISGRLLFNLM	120						
Db	61	ILQIRHTPHNGLSDIYLLNDPHEAFAARAALIESAHSLSLDQYIWRNDISGRLLFNLV	120						
Qy	121	YLAAGRGVRVRLDDNNTRGLDLDLLALDSHPNIEVRLNFPFVLRKWRALGYLTDPPRL	180						
Db	121	YLAAGRGVRVRLDDNNTRGLDLDLLALDSHPNIEVRLNFPFVLRKWRALGYLTDPPRL	180						
Qy	181	NRRMKNKSFADNRAITILGGRNIGDEYFKVGEDTVFADLDLATGSVVGSEVSHDFRYWA	240						
Db	181	NRRMKNKSFADNRAITILGGRNIGDEYFKVGEDTVFADLDLATGSVVGSEVSHDFRYWA	240						
Qy	241	SHSAHNATRIIRSGNIGKGLQALGYNDETSRHALLRYRETVESQSPLYQKIQTGRIDWQSV	300						
Db	241	SHSAHNATRIIRSGNIGKGLQALGYNDETSRHALLRYRETVESQSPLYQKIQTGRIDWQSV	300						
Qy	301	QTRLSIDTTPAKGLDRDRKPPITAGRLQDALQPEKSVLYSPYFVPTKSGTDALAKLVQD	360						
Db	301	QTRLSIDTTPAKGLDRDRKPPITAGRLQDALQPEKSVLYSPYFVPTKSGTDALAKLVQD	360						
Qy	361	GIDVTVLTNSLQATDVAHVHSGYKVRKPLLKAGIKLYELQPNHAVPATKDKGLTGSSVT	420						
Db	361	GIDVTVLTNSLQATDVAHVHSGYKVRKPLLKAGIKLYELQPNHAVPATKDKGLTGSSVT	420						

cardiolipin syntha
cardiolipin syntha
cardiolipin syntha
cardiolipin synthe
probable phospholi
cardiolipin syntha
cardiolipin syntha
cardiolipin syntha
probable phospholi
yhbO protein - Esc
probable synthetas
cardiolipin synthe
cardiolipin synthe
cardiolipin syntha
cardiolipin syntha
cardiolipin syntha

QY 421 SLHAKTFIVDGKRIFIGSFNLDPRGARNLTENGWVIESPKIAEQMERTLADTTPYAYRV 480
DB 421 SLHAKTFIVDGKRIFIGSFNLDPRGARNLTENGWVIESPKIAEQMERTLADTSPYAYRV 480
QY 481 TLDKHNRQLQWHDPAATKYPNEPEAKLWKRIAATKILSLPIEGLL 525
DB 481 TLDNRNLQWHDPAATKYPNEPEAKLWKRIAATKILSLPIESLL 525
RESULT 2
B81083
cardiolipin synthetase family protein NMB1434 [imported] - Neisseria meningitidis (strain B81083)
C:Species: Neisseria meningitidis
C>Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 09-Jul-2004
C:Accession: B81083
R:Testelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, J.A.; Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.A.; Ri, H.; Qin, H.; Vamathevan, J.; Gill, J.; Scariato, V.; Masignani, V.; Pizza, M. Science 287, 1809-1815, 2000
A:Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.; Vercellotti, R. Nature 403, 645-649, 2000
A:Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MC58.
A:Reference number: A81000; MUID:20175755; PMID:10710307
A:Accession: B81083
A:Status: preliminary
A:Gene: NMB1434
A:Molecule type: DNA
A:Residues: 1-508 <TET>
A:Cross-references: UNIPROT:Q9JYU0; UNIPARC:UPI00000C46B9; GB:AE002494; GB:AE002098; NID:10710307
A:Experimental source: serogroup B, strain MC58
C:Genetics:
C:Superfamily: cardiolipin synthase
Query Match 94.6%; Score 2573; DB 2; Length 508;
Best Local Similarity 97.8%; Pred. No. 5.2e-177;
Matches 497; Conservative 5; Mismatches 6; Indels 0; Gaps 0;
QY 18 MKTRSLISLLCLLSCSWLPPLEERTESRHFNTPSKVRLDNLIIQIRHTPHNGLSDIY 77
DB 1 MKTRSLISLLCLLSCSWLPPLEERTESRHFNTPSKVRLDNLIIQIRHTPHNGLSDIY 60
QY 78 LDDPHEALARAALIESAHSLSLDQYIWRNDISGRLLFNLMYLAARGVRLDLDN 137
DB 61 LNDPHEAFARAALIESAHSLSLDQYIWRNDISGRLLFNLMYLAARGVRLDLDN 120
QY 138 NTRGLDDLLLALDHPNIEVRLFNPFLRKWRALCYLTDFFPLNRRMNKSFADNRATI 197
DB 121 NTRGLDDLLLALDHPNIEVRLFNPFLRKWRALCYLTDFFPLNRRMNKSFADNRATI 180
QY 198 LGGRNIGDEYFKVGEDTFADLIDLATGVSVEVSHDFDRYWAHSAHNATRIIRSGNIG 257
DB 181 LGGRNIGDEYFKVGEDTFADLIDLATGVSVEVSHDFDRYWAHSAHNATRIIRSGDIG 240
QY 258 KGLQALGYNDETSRALLRYRETVEOSPLYQKIQTGRIDWQSVQTRLSIDTPAKGLDRDR 317
DB 241 KGLQALGYNDETSRALLRYRETVEOSPLYQKIQTGRIDWQSVQTRLSIDTPAKGLDRDR 300
QY 318 RKPTIAGRLQDALKQPEKSVYLVSPFYVPTKSGTDLAKLVQDGDIVDTVLTNSLQATDVA 377
DB 301 RKPTIAGRLQDALKQPEKSVYLVSPFYVPTKSGTDLAKLVQDGDIVDTVLTNSLQATDVA 360
QY 378 AVHSGYVKYRPELLKAGIKLYELQPNHAPATKDKGLTGSSVTSIHAKTFFIVDGKRIEIG 437
DB 361 AVHSGYVKYRPELLKAGIKLYELQPNHAPATKDKGLTGSSVTSIHAKTFFIVDGKRIEIG 420
QY 438 SFNLDPGRARNLTENGWVIESPKIAEQMERTLADTTPYAYRVTLDKHNRQLQWHDPAATRK 497
DB 421 SFNLDPGRARNLTENGWVIESPKIAEQMERTLADTTPYAYRVTLDRNRNLQWHDPAATRK 480
QY 498 TYPNEPEAKLWKRIAATKILSLPIEGLL 525
DB 481 TYPNEPEAKLWKRIAATKILSLPIEGLL 508

RESULT 3
C64847
ymdC protein - Escherichia coli (strain K-12)
C:Species: Escherichia coli
C>Date: 12-Sep-1997 #sequence_revision 17-Sep-1997 #text_change 09-Jul-2004
C:Accession: C64847
R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Co.
A.; Rose, D.J.; Mau, B.; Shao, Y. Science 277, 1453-1462, 1997
A:Title: The complete genome sequence of Escherichia coli K-12.
A:Reference number: A64720; MUID:97426617; PMID:9278503
A:Accession: C64847
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-493 <BLAT>
A:Cross-references: UNIPROT:P75919; UNIPARC:UPI000013B993; GB:AE000206; GB:U00096; NID:9
A:Experimental source: strain K-12, substrain MG1655
C:Genetics:
A:Gene: ymdC
C:Superfamily: cardiolipin synthase
Query Match 40.5%; Score 1102.5; DB 2; Length 493;
Best Local Similarity 47.6%; Pred. No. 2.2e-71;
Matches 221; Conservative 83; Mismatches 143; Indels 17; Gaps 5;
QY 72 GLSDIYLLDDPHEALARAALIESAHSLSLDQYIWRNDISGRLLFNLMYLAARGVVR 131
DB 37 QCCGLFPLEKSLDAFAARYRLAEMAHTLDVQYVIWQDDMSGRLLFSAALLAAKRGVR 96
QY 132 LLLDDNTRGLDLLLLALDHPNIEVRLFNPFLRKWRALCYLTDFFPLNRRMNKSFATA 191
DB 97 LLLDDNTRPGDDDLRLDLDHPNIEVRLFNPFLRLRPLGYITDFSLNRRMNKSFIV 156
QY 192 DNRTATILGGRNIGDEYFKVGEDTFADLIDLATGVSVEVSHDFDRYWAHSAHNATRII 251
DB 157 DGVTVLGGRNIGDAYFCAGEEPLFSDLDWAIGPVVEDVADDFARYWYCKSVPLQV 216
QY 252 RSGNIGK-----LQALGYNDETSRALLRYRETVEOSPLYQKIQTGRIDWQSVQTRL 304
DB 217 ---DVPEGMADRIELPASWHNDAMTH----RYLRKMESSPFINHLVDGTPLIWAKTRL 269
QY 305 ISDTPAKGLDRDRKPPITAGRLQDALKQPEKSVYLVSPFYVPTKSGTDLAKLVQDGDIV 364
DB 270 LSDDPKAGEKAKRHSLLPQRLFDIMGSPSERIDIISSYFVPTTAGVLAQLLRWVRKGVKI 329
QY 365 TVLTNSLQATDVAAVHSGYVKYRPELLKAGIKLYELQPNHAPAT-KDKGLTGSSVTSLSH 423
DB 330 AILLNSLAANDVAVVHAGYARWRKLLRYGVVELYELKPTREQSSTLHDRGITGNSGASLH 389
QY 424 AKTIVDGKRIFGSFNLDPRSARLNTMGVVIESPKIAEQMERTLADTTPYAYRVTL 483
DB 390 AKTFSIDGKTVFIGSFNLDPRSTLLNTEMGFVIESETLQALIDRFRFIQSDYDAWQLRLD 449
QY 484 KHNRQLQWHDPAATK-TYPNEPEAKLWKRIAATKILSLPIEGLL 525
DB 450 RWGRINWDRHAKKEIILKKEPATSFWRVVRVLRVLAASILPVEWLL 493
RESULT 4
H90806
probable synthase [imported] - Escherichia coli (strain O157:H7, substrain RIMD 0509952)
C:Species: Escherichia coli
C>Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 09-Jul-2004
C:Accession: H90806
R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H. DNA Res. 8, 11-22, 2001
A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and geno
A:Reference number: A99629; MUID:21156231; PMID:11258796
A:Accession: H90806
A:Status: preliminary
A:Molecule type: DNA

A;Residues: 1-493 <HAY>
A;Cross-references: UNIPROT:Q8X9I7; UNIPARC:UPI00000D07FD; GB:BA000007; PIDN:BA034847.1;
A;Experimental source: strain O157:H7, substrain RIMD 050952
C;Genetics:
A;Gene: ECs1424
C;Superfamily: cardiolipin synthase

Query Match 40.5%; Score 1100.5; DB 2; Length 493;
Best Local Similarity 47.4%; Pred. No. 3e-71;
Matches 220; Conservative 84; Mismatches 143; Indels 17; Gaps 5;

```
QY 72 GLSDIYLLDDPHEALAAALIESAEHSLDQYIWRNDISGRLLFNLMYLAERGVRVR 131
Db |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
37 GQGLFPLEKSLDAFAARYLAEMSEHTLDVQYIWDMSGRLLFSALLAAAKRGVRVR 96
QY 132 LLLDDNNTRGLDDLLALDHPNIEVLFNFPVLRKWRALCYLTDFFPLNRRMHNKSF 191
Db |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
97 LLLDDNNTPGLDDILRLDSDHPRIEVLRFNFPFSLRLLRPLGYITDFSLNRRMHNKSF 156
QY 192 DNRATILGGRNIGDEYFKVGEDTVFADLDILATGSVVGESHDPDRYWAHSHAHNATRII 251
Db |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
157 DGVVTLVGGRNIGDAYFGAGEEPLFSDLDVMAIGPVVEDVADDFARYWYCKSVSPLOQVL 216
QY 252 RSGNIGKG-----LQALGVNDETSRHALLRYRETVEQSPLYOKIQTGRIDWOSVOTRL 304
Db |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
217 ---DVPEGMADRIELPASHNDAMTH-----RYLRKMESPFFINHLVDGTLPLIWAQRL 269
QY 305 ISDTPAKGLDRDRKPPKIAGRLQDALQKPEKSVVLVSPFYFPTKSGTDALAKLVQDGDIV 364
Db |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
270 LSDDPAKGEKAKHSLLPQRLFDIMGSPSERDIISSYFVPTAGVAQLLRWVRKVKVI 329
QY 365 TVLTNSLQATDVAHVSHGYKRYKPLKAGIKYELQPNHAVPAT-KDKGLTGSSVTSLSH 423
Db |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
330 AILTNSLAANDVAVVHAGYARWRKLLRYGVELYELKPTREQSTLHDRGITGNSGASLH 389
QY 424 AKTFIVDGKRIFTGSFNLDRPSARLNTMGVIVIESPKIAEQMERTLADTTPEYAYRVTL 483
Db |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
390 AKTFSIDGKTVFGSFNFDPRSTLLNTMGFVIESETLAQLIDKRFQSQDYDAAWQLRLD 449
QY 484 KHNRLQWHDPATRK--TYPNEPEAKLWKRIAATKLSLLPIEGLL 525
Db |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
450 RWGRINWVDHRHAKKEIVLKEPATSFWKRVWVRILASILPVEWLL 493
```

RESULT 5
D85666
probable synthase ymdC [imported] - Escherichia coli (strain O157:H7, substrain EDL933)
C;Species: Escherichia coli
C;Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004
C;Accession: D85666
R;Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,
Nature 409, 529-533, 2001
A;Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A;Reference number: A85480; MUID:21074935; PMID:11206551
A;Accession: D85666
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-493 <STO>
A;Cross-references: UNIPROT:Q8X9I7; UNIPARC:UPI0000165766; GB:AE005174; NID:g12514578; F
A;Experimental source: strain O157:H7, substrain EDL933
C;Genetics:
A;Gene: ymdC
C;Superfamily: cardiolipin synthase

Query Match 40.2%; Score 1093.5; DB 2; Length 493;
Best Local Similarity 47.2%; Pred. No. 9.6e-71;
Matches 219; Conservative 84; Mismatches 144; Indels 17; Gaps 5;

```
QY 72 GLSDIYLLDDPHEALAAALIESAEHSLDQYIWRNDISGRLLFNLMYLAERGVRVR 131
Db |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
37 GQGLFPLEKSLDAFAARYLAEMSEHTLDVQYIWDMSGRLLFSALLAAAKRGVRVR 96
```

```
QY 132 LLLDDNNTRGLDDLLALDHPNIEVLFNFPVLRKWRALCYLTDFFPLNRRMHNKSF 191
Db |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
97 LLLDDNNTPGLDDILRLDSDHPRIEVLRFNFPFSLRLLRPLGYITDFSLNRRMHNKSF 156
QY 192 DNRATILGGRNIGDEYFKVGEDTVFADLDILATGSVVGESHDPDRYWAHSHAHNATRII 251
Db |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
157 DGVVTLVGGRNIGDAYFGAGEEPLFSDLDVMAIGPVVEDVADDFARYWYCKSVSPLOQVL 216
QY 252 RSGNIGKG-----LQALGVNDETSRHALLRYRETVEQSPLYOKIQTGRIDWOSVOTRL 304
Db |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
217 ---DVPEGMADRIELPASHNDAMTH-----RYLRKMESPFFINHLVDGTLPLIWAQRL 269
QY 305 ISDTPAKGLDRDRKPPKIAGRLQDALQKPEKSVVLVSPFYFPTKSGTDALAKLVQDGDIV 364
Db |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
270 LSDDPAKGEKAKHSLLPQRLFDIMGSPSERDIISSYFVPTAGVAQLLRWVRKVKVI 329
QY 365 TVLTNSLQATDVAHVSHGYKRYKPLKAGIKYELQPNHAVPAT-KDKGLTGSSVTSLSH 423
Db |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
330 AILTNSLAANDVAVVHAGYARWRKLLRYGVELYELKPTREQSTLHDRGITGNSGASLH 389
QY 424 AKTFIVDGKRIFTGSFNLDRPSARLNTMGVIVIESPKIAEQMERTLADTTPEYAYRVTL 483
Db |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
390 AKTFSIDGKTVFGSFNFDPRSTLLNTMGFVIESETLAQLIDKRFQSQDYDAAWQLRLD 449
QY 484 KHNRLQWHDPATRK--TYPNEPEAKLWKRIAATKLSLLPIEGLL 525
Db |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
450 RWGRINWVDHRHAKKEIVLKEPATSFWKRVWVRILASILPVEWLL 493
```

RESULT 6

AE0636

conserved hypothetical protein STY1185 [imported] - Salmonella enterica subsp. enterica

C;Species: Salmonella enterica subsp. enterica serovar Typhi

A;Note: this species has also been called Salmonella typhi

C;Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 09-May-2004

C;Accession: AE0636

R;Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Main, J.; Churcher

th, T.; Connerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar

S.; Moulie, S.; O'Gaora, P.

Nature 413, 848-852, 2001

A;Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.

A;Title: Complete genome sequence of a multiple drug resistant Salmonella enterica sero

A;Reference number: AB0502; MUID:21534947; PMID:11677608

A;Accession: AE0636

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-495 <PAR>

A;Cross-references: UNIPARC:UPI0000059FC8; GB:AL513382; PIDN:CAD08272.1; PMID:g16502319;

C;Genetics:

A;Gene: STY1185

C;Superfamily: cardiolipin synthase

Query Match 40.1%; Score 1089.5; DB 2; Length 495;

Best Local Similarity 47.4%; Pred. No. 1.9e-70;

Matches 221; Conservative 77; Mismatches 147; Indels 21; Gaps 5;

```
QY 72 GLSDIYLLDDPHEALAAALIESAEHSLDQYIWRNDISGRLLFNLMYLAERGVRVR 131
Db |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
39 GEGCILALDNLSDAFAARYLRTMAAARTLDVQYIWEEDMSGRLLFSALLAAAKRGVHR 98
QY 132 LLLDDNNTRGLDDLLALDHPNIEVLFNFPVLRKWRALCYLTDFFPLNRRMHNKSF 191
Db |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
99 LLLDDNNTPGLDDTLRLDSDHPNIEVLFNFPFSLRLLRPLGYITDFARLNRHNNKSYTA 158
QY 192 DNRATILGGRNIGDEYFKVGEDTVFADLDILATGSVVGESHDPDRYWAHSHAHNATRII 251
Db |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
159 DGVVTLVGGRNIGDAYFGAGEEPLFSDLDVMAIGPVVNDVANDFERWYRWCSSVSTLQOVL 218
QY 252 RSGNIGKGLQAL-----GVNDETSRHALLRYRETVEQSPLYOKIQTGRIDWOSVOT 302
Db |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
219 SLSE-----QELTQRIELPESWYNDEITR-----RYLHKLETQSFMAIDLDRGTLPLIWAQ 269
QY 303 RLISDTPAKGLDRDRKPPKIAGRLQDALQKPEKSVVLVSPFYFPTKSGTDALAKLVQDGI 362
```

```

Db 270 RLLSDRSKGEKQARHSLLPQRLFDMVSGSTERIDIISAYFVTRAGVAQLNLVRKGV 329
Qy 363 DVTVLTNSLOATDVAAVHSGYVKYKPKLLKAGIKLYELOPNHA-VPA TKQKGLTGSSVTS 421
Db 330 KIAILTNSLAANDVAHVHAGYARWRKLLRYGVBELKPTREHETAVHDRGLTNGSGS 389
Qy 422 LKAKTFVDGKRIPIGPNLDRSARLNTENGVIIESPKIAEQMERLADTTPYAYRVT 481
Db 390 LKAKTFSDGSKVIGSLNDFPRSTLNTENGFIETSLATLTKRFTSQORDAAWOLR 449
Qy 482 LDKHNRLQWHD--PATRKYTNEPEAKLWKRIAAKILSLPIEGLL 525
Db 450 LDRNGRLNWDROQBEKVKKEPATRFQWRVLRVLAAILPVEWLL 495

RESULT 7
G97556
hypothetical 55_9K protein in csgc-mdog intergenic region [imported] - Agrobacterium tum
C:Species: Agrobacterium tumefaciens
C:Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 09-Jul-2004
C:Accession: G97556
R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Quroollo, B.; Goldman,
A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.;
Science 294, 2323-2328, 2001
A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum
A:Reference number: A97359; MUID:21608551; PMID:11743194
A:Accession: G97556
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-518 <KUR>
A:Cross-references: UNIPROT:Q8UEX3; UNIPARC:UPI00000D1C0B; GB:AE007869; PIDN:AAK87408.1;
C:Genetics:
A:Map position: circular chromosome
A:Superfamily: cardiolipin synthase

Query Match 29.7%; Score 807; DB 2; Length 518;
Best Local Similarity 38.6%; Pred. No. 4e-50;
Matches 204; Conservative 85; Mismatches 196; Indels 44; Gaps 16;

Qy 23 LISLCLLLCSCSWLPPLERTESRHFNTSKPVLLDNIQLIRTPH--NNGLSD---IYL 78
Db 8 IIIILLMIGPSLFVVGKQREKAIPKRPSTALPTEDETALDRHQWIRNGWNEKNALCL 67
Qy 79 LDDPHEALAAARALIESAEHSLDQYIWRNDISGRLLFNLMYLAARGVVRVRLLLDDNN 138
Db 68 LHSNLDAPAVRVAARAAGRSGLDMYYMWNADLTGRLMMREVIATAADRGVVRVRLLLDDLG 127
Qy 139 TRGLDDLLALLDHPNIEVRLFNPFVLRK--WRALGYLTDPPRLNRRMHNKSFADNRA 195
Db 128 VMSMDRIFPHADSHPNIELRFLNPTREANILHRSLELVLFPRSVNRRMHNKAWIADGRA 187
Qy 196 TILGGRNIGDEYFKVGEDTVFADLDLATGSGVGEVSHDFRYWASHAHN-----AT 248
Db 246 RPNKLAKLRRELDALPOSE-----AARPYLERVEEQ--YGRDHFLMSDRLHW--VDTADV 296
Qy 305 ISDTPAKGLDRDRKPP--IAGRLQDALKQPEKSVLYVSPYFVPTKSGTDALAKLVQDGI 362
Db 297 LADPPEKAAGK-RRKGNHFLMESLLPLMQAAGESLHITSFYPIPKQGVFEIFDLAERGV 355
Qy 363 DVTVLTNSLOATDVAAVHSGYVKYKPKLLKAGIKLYELOPNHNAVPA TKDKG---LTGSSV 419
Db 246 RPNKLAKLRRELDALPOSE-----AARPYLERVEEQ--YGRDHFLMSDRLHW--VDTADV 296
Qy 305 ISDTPAKGLDRDRKPP--IAGRLQDALKQPEKSVLYVSPYFVPTKSGTDALAKLVQDGI 362
Db 297 LADPPEKAAGK-RRKGNHFLMESLLPLMQAAGESLHITSFYPIPKQGVFEIFDLAERGV 355
Qy 363 DVTVLTNSLOATDVAAVHSGYVKYKPKLLKAGIKLYELOPNHNAVPA TKDKG---LTGSSV 419
Db 356 SLAILTNSLAATDVAHVHAGYARWRKLLSGVRLHELR-----SQADQGSFTLRGSGQ 409
Qy 420 TSLHAKTFVDGKRIPIGSNLDRPSARLNTENGVIIESPKIAEQMERLADTTPYAY 478
Db 410 ASLHTKAFTRDGETGYIGSLNDFPRSLNTEMGVFNPSAPLVARMDEIFAEEIRRTMSF 469

Query Match 29.7%; Score 807; DB 2; Length 518;
Best Local Similarity 38.6%; Pred. No. 4e-50;
Matches 204; Conservative 85; Mismatches 196; Indels 44; Gaps 16;

Qy 23 LISLCLLLCSCSWLPPLERTESRHFNTSKPVLLDNIQLIRTPH--NNGLSD---IYL 78
Db 8 IIIILLMIGPSLFVVGKQREKAIPKRPSTALPTEDETALDRHQWIRNGWNEKNALCL 67
Qy 79 LDDPHEALAAARALIESAEHSLDQYIWRNDISGRLLFNLMYLAARGVVRVRLLLDDNN 138
Db 68 LHSNLDAPAVRVAARAAGRSGLDMYYMWNADLTGRLMMREVIATAADRGVVRVRLLLDDLG 127
Qy 139 TRGLDDLLALLDHPNIEVRLFNPFVLRK--WRALGYLTDPPRLNRRMHNKSFADNRA 195
Db 128 VMSMDRIFPHADSHPNIELRFLNPTREANILHRSLELVLFPRSVNRRMHNKAWIADGRA 187
Qy 196 TILGGRNIGDEYFKVGEDTVFADLDLATGSGVGEVSHDFRYWASHAHN-----AT 248
Db 246 RPNKLAKLRRELDALPOSE-----AARPYLERVEEQ--YGRDHFLMSDRLHW--VDTADV 296
Qy 305 ISDTPAKGLDRDRKPP--IAGRLQDALKQPEKSVLYVSPYFVPTKSGTDALAKLVQDGI 362
Db 297 LADPPEKAAGK-RRKGNHFLMESLLPLMQAAGESLHITSFYPIPKQGVFEIFDLAERGV 355
Qy 363 DVTVLTNSLOATDVAAVHSGYVKYKPKLLKAGIKLYELOPNHNAVPA TKDKG---LTGSSV 419
Db 356 SLAILTNSLAATDVAHVHAGYARWRKLLSGVRLHELR-----SQADQGSFTLRGSGQ 409
Qy 420 TSLHAKTFVDGKRIPIGSNLDRPSARLNTENGVIIESPKIAEQMERLADTTPYAY 478
Db 410 ASLHTKAFTRDGETGYIGSLNDFPRSLNTEMGVFNPSAPLVARMDEIFAEEIRRTMSF 469
```

```

Qy 479 RVTLDKHNRLQW--HDPATRKTYNPEPEAKLWKRIAAKILSLPIEGLL 525
Db 470 ELDIDSANRIVWMTTEERGQPKIYRREPDAAISRIIAGIMRWVLPLESOL 518

RESULT 8
AB2777
Phospholipase D family protein [imported] - Agrobacterium tumefaciens (strain C58, Dupon
C:Species: Agrobacterium tumefaciens
C:Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 09-Jul-2004
C:Accession: AB2777
R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, L.
erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutayavin, T.; Levy, R.; Li, M.; McClell
; Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,
ster, E.W.
A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A:Reference number: AB2577; MUID:21608550; PMID:11743193
A:Accession: AB2777
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-518 <KUR>
A:Cross-references: UNIPROT:Q8UEX3; UNIPARC:UPI00000D1C0B; GB:AE008688; PIDN:AAL42632.1;
A:Experimental source: strain C58 (Dupont)
C:Genetics:
A:Gene: Atul630
A:Map position: circular chromosome
C:Superfamily: cardiolipin synthase

Query Match 29.7%; Score 807; DB 2; Length 518;
Best Local Similarity 38.6%; Pred. No. 4e-50;
Matches 204; Conservative 85; Mismatches 196; Indels 44; Gaps 16;

Qy 23 LISLCLLLCSCSWLPPLERTESRHFNTSKPVLLDNIQLIRTPH--NNGLSD---IYL 78
Db 8 IIIILLMIGPSLFVVGKQREKAIPKRPSTALPTEDETALDRHQWIRNGWNEKNALCL 67
Qy 79 LDDPHEALAAARALIESAEHSLDQYIWRNDISGRLLFNLMYLAARGVVRVRLLLDDNN 138
Db 68 LHSNLDAPAVRVAARAAGRSGLDMYYMWNADLTGRLMMREVIATAADRGVVRVRLLLDDLG 127
Qy 139 TRGLDDLLALLDHPNIEVRLFNPFVLRK--WRALGYLTDPPRLNRRMHNKSFADNRA 195
Db 128 VMSMDRIFPHADSHPNIELRFLNPTREANILHRSLELVLFPRSVNRRMHNKAWIADGRA 187
Qy 196 TILGGRNIGDEYFKVGEDTVFADLDLATGSGVGEVSHDFRYWASHAHN-----AT 248
Db 188 VIVGGRNIGDAYFDAERANFHDIFLGFGIVADATEIFDDYW--NSAVSPVRSLLAR 245
Qy 249 RIIRSGNIGKGLQALGYNDETSRHALLRYRETVEQSPLYQK---IQGRIDWQSVQTR-L 304
Db 246 RPNKLAKLRRELDALPOSE-----AARPYLERVEEQ--YGRDHFLMSDRLHW--VDTADV 296
Qy 305 ISDTPAKGLDRDRKPP--IAGRLQDALKQPEKSVLYVSPYFVPTKSGTDALAKLVQDGI 362
Db 297 LADPPEKAAGK-RRKGNHFLMESLLPLMQAAGESLHITSFYPIPKQGVFEIFDLAERGV 355
Qy 363 DVTVLTNSLOATDVAAVHSGYVKYKPKLLKAGIKLYELOPNHNAVPA TKDKG---LTGSSV 419
Db 356 SLAILTNSLAATDVAHVHAGYARWRKLLSGVRLHELR-----SQADQGSFTLRGSGQ 409
Qy 420 TSLHAKTFVDGKRIPIGSNLDRPSARLNTENGVIIESPKIAEQMERLADTTPYAY 478
Db 410 ASLHTKAFTRDGETGYIGSLNDFPRSLNTEMGVFNPSAPLVARMDEIFAEEIRRTMSF 469

Query Match 29.7%; Score 807; DB 2; Length 518;
Best Local Similarity 38.6%; Pred. No. 4e-50;
Matches 204; Conservative 85; Mismatches 196; Indels 44; Gaps 16;

Qy 23 LISLCLLLCSCSWLPPLERTESRHFNTSKPVLLDNIQLIRTPH--NNGLSD---IYL 78
Db 8 IIIILLMIGPSLFVVGKQREKAIPKRPSTALPTEDETALDRHQWIRNGWNEKNALCL 67
Qy 79 LDDPHEALAAARALIESAEHSLDQYIWRNDISGRLLFNLMYLAARGVVRVRLLLDDNN 138
Db 68 LHSNLDAPAVRVAARAAGRSGLDMYYMWNADLTGRLMMREVIATAADRGVVRVRLLLDDLG 127
Qy 139 TRGLDDLLALLDHPNIEVRLFNPFVLRK--WRALGYLTDPPRLNRRMHNKSFADNRA 195
Db 128 VMSMDRIFPHADSHPNIELRFLNPTREANILHRSLELVLFPRSVNRRMHNKAWIADGRA 187
Qy 196 TILGGRNIGDEYFKVGEDTVFADLDLATGSGVGEVSHDFRYWASHAHN-----AT 248
Db 188 VIVGGRNIGDAYFDAERANFHDIFLGFGIVADATEIFDDYW--NSAVSPVRSLLAR 245
Qy 249 RIIRSGNIGKGLQALGYNDETSRHALLRYRETVEQSPLYQK---IQGRIDWQSVQTR-L 304
Db 246 RPNKLAKLRRELDALPOSE-----AARPYLERVEEQ--YGRDHFLMSDRLHW--VDTADV 296
Qy 305 ISDTPAKGLDRDRKPP--IAGRLQDALKQPEKSVLYVSPYFVPTKSGTDALAKLVQDGI 362
Db 297 LADPPEKAAGK-RRKGNHFLMESLLPLMQAAGESLHITSFYPIPKQGVFEIFDLAERGV 355
Qy 363 DVTVLTNSLOATDVAAVHSGYVKYKPKLLKAGIKLYELOPNHNAVPA TKDKG---LTGSSV 419
Db 356 SLAILTNSLAATDVAHVHAGYARWRKLLSGVRLHELR-----SQADQGSFTLRGSGQ 409
Qy 420 TSLHAKTFVDGKRIPIGSNLDRPSARLNTENGVIIESPKIAEQMERLADTTPYAY 478
Db 410 ASLHTKAFTRDGETGYIGSLNDFPRSLNTEMGVFNPSAPLVARMDEIFAEEIRRTMSF 469

RESULT 9
F82983
```

conserved hypothetical protein PA5310 [imported] - Pseudomonas aeruginosa (strain PA01)
C:Species: Pseudomonas aeruginosa
C>Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2004
C:Accession: F82983
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Muzoguchi, S.D.; Warren, P.; Hickey, M.J.; Br
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,
.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho
gen
A:Reference number: A82950; MUID:20437337; PMID:10384043
A:Accession: F82983
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-529 <STO>
A:Cross-references: UNIPROT:Q9HTP4; UNIPARC:UPI00000C5F9D; GB:AE004943; GB:AE004091; NID
A:Experimental source: strain PA01
C:Genetics:
A:Gene: PA5310
C:Superfamily: cardiolipin synthase
Query Match 29.1%; Score 792; DB 2; Length 529;
Best Local Similarity 36.7%; Pred. No. 4.9e-49;
Matches 200; Conservative 86; Mismatches 189; Indels 70; Gaps 16;
QY 21 RSISLLCLLCCSSWLPP-----LEERTESHFNSTKPEVLNDNLIQRHTP 68
DB 15 RRFLLALLALSGCASTPPQPSSALPAEGTWLARQAE-----IQGRDHP 59
QY 69 HNNGLSDIYLLDPHEALARAALIESAEHSLDIOYIWRNDISGRLLFNMLYLAERG 128
DB 60 ---GQSGFHLLSASEDAFARAALIRAAQSLDIOYIYVHDGLTTALAYELLKAADRGV 116
QY 129 RVRLLDNNTRGLDLLLALDGHNPENVEFLNFPVLRK---WRALGYLTDPPRLNRM 184
DB 117 RVRLIDDTASDGDWDEIGVLSHPNLOVRLNPLHLGRATGTRGVGRFLNLSQQHRRM 176
QY 185 HNKSFADNRATILGGNIGDEYFKVGEDTVFADLDILATGVSVVGESHDFRYWASHA 244
DB 177 HNKWLADGTAIVGGNLDGEYFNAPENMFTDLDLLGVPIANQLSHSFDQYVNSAIS 236
QY 245 HNAT---RIIRSGNIGKGLQALGYNDSTSRHLLY---RETVEQS-----PLYQKIQTR 294
DB 237 RPIDFLWRAPYGEEL-----ESARRKLQRYLRKESKESGYTRHFLDRGDQPR 285
QY 295 I-DWQS---VQTRLISDTPAKGLDRDRKPP--IAGRLQDALKQEKSVLYVSPYFVPT 347
DB 286 LGNWLNLNLTWARAEATWADPKVLSRGEPPHLLSPHLAAGLFGVQKEILLVSAIFVPA 345
QY 348 KSGTDALAKLVODGIDVTVLTNSLOATDVAAHSGYVKYRKPILLKAGIKLYELQPNHVP 407
DB 346 KGLNLYLTGKADSGVRVRLTNSLEATDVPVHAGVAPYRMALLEHGVKLYELRANPDQP 405
QY 408 AT-KDKGLTSGSVTSLHAKTFIVDGKRIFFGSLNLDPRSRARLNTMGVIESPKIAEQME 466
DB 406 LSGAPWRLHGSASSLHSAKAWPDRKRVFTGSGFDPFRSLMNTVEGVIVDPSPLAEQVR 465
QY 467 R-TLADTTPYAYRVTLDKHN---RLQWHD--PATKTYNPEPEAKLWKRIAKILSLP 520
DB 466 QLALEGMAVSQVRIIDRSGSRPKLVWIDRGRQAQVLNHEP-GSLWRRLNNAWAGMIG 524
QY 521 IEGLL 525
DB 525 LEKML 529
RESULT 10
E71963
probable cardiolipin synthase - Helicobacter pylori (strain J99)
C:Species: Helicobacter pylori
A:Variety: strain J99
C>Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 09-Jul-2004
C:Accession: E71963
R:Alm, R.A.; Ling, L.S.L.; Moir, D.T.; King, B.L.; Brown, E.D.; Doig, P.C.; Smith, D.R.;

; Ives, C.; Gibson, R.; Merberg, D.; Mills, S.D.; Jiang, Q.; Taylor, D.E.; Vovis, G.F.;
Nature 397, 176-180, 1999
A:Title: Genomic sequence comparison of two unrelated isolates of the human gastric pat
hogen
A:Reference number: A71800; MUID:99120557; PMID:9923682
A:Accession: E71963
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-502 <ARN>
A:Cross-references: UNIPROT:Q9ZMP2; UNIPARC:UPI000013942E; GB:AE001456; GB:AE001439; NI
A:Experimental source: strain J99
C:Genetics:
A:Gene: jhp0176
C:Superfamily: cardiolipin synthase
Query Match 23.2%; Score 632; DB 2; Length 502;
Best Local Similarity 32.8%; Pred. No. 1.4e-37;
Matches 161; Conservative 93; Mismatches 181; Indels 56; Gaps 17;
QY 64 IRHTPHNNGLSDIY-----LLDDPHEALARAALIESAEHSLDLQYYIWRN 109
DB 30 ISYDPTTTIGSLYAKNLKENPKHSAAILLEDGFDALLHRVGLIRMSQKSIDMQTYIYKN 89
QY 110 DISGRLLFNMLYLAERGVRVRLLDNNTNTRGLD----DLLLALDHPNTEVRLFNPFVL 165
DB 90 DLSQVIAKELLNANRGVKVRIILDDN---GLDSDFSDIML-LNFHKNIEVKIFNPYYI 145
QY 166 RKWRALGY---LTDFFPLARRMHNKSFADNRATILGGNIGDEYFKVGEDTVFADLDIL 222
DB 146 RN-KGURYFEMLADYERIKRMHNKLFIVDNFAVIIGGRNIGYFDNDLDTNPLDLDAL 204
QY 223 ATGSVVGVESHDPDRYVASHNAHNATRIIRS-----GNICKGLQALGYNDSTSRHAL 274
DB 205 FFGVASKAKESPENYWRFRHSIPVS-LURTHKLNKNVKEIAKLEKIPISAEADANEFE 263
QY 275 LRYRETVESPLYQ-KIQTRIDWQSVQTRLISDTPAKGLDRDRRKPPIAGRIQDALKQP 333
DB 264 KKVNDPIERFQKYQYPIYVG-----NAIFLADLPK-IDTPLY-S-PKIAFEKALKNA 314
QY 334 EKSVLYVSPYFVPTKSGTDALAKLVODGIDVTVLTNSLOATDVAAHSGYVKYRKPILLKA 393
DB 315 KDSVFASISYFPGKKIMKIFKNQISKGIELNLTNSLSDTAIVVYGAWERYNRKLVRM 374
QY 394 GKLYELQPNHVAHPATKDKGLTSGSVTSLHAKTFIVDGKRIFFGSLNLDPRSRARLNTMG 453
DB 375 GANVYIERNDFNRQIKGR---FSTGKLSLHGKTIIVFDALTLLGSGNIDPRSAINTESA 431
QY 454 VIESPKIAEQMERTIADTTPYAYRVTLDKNRLQWHDPAKTKTY-----PNPEAKLWK 509
DB 432 VLFDNPSFAKRVRLSLKHA-QQSWHLVLYRH-RVIWE--ATEGILIHEKNSPDTSF 487
QY 510 RIAAKILSLP 520
DB 488 RLKIKESKVL 498
RESULT 11
F64543
conserved hypothetical secreted protein HP0190 - Helicobacter pylori (strain 26695)
C:Species: Helicobacter pylori
C>Date: 09-Aug-1997 #sequence_revision 09-Aug-1997 #text_change 09-Jul-2004
C:Accession: F64543
R:Tomb, J.F.; White, O.; Kerlavage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R.D
Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khalak, H.G.; Glodek, A.; McKen
son, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.; Bowman, C.; Watthey, I.
Nature 386, 539-547, 1997
A:Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karpk, P.D.; Smith, H.O.; Fraser, C
A:Title: The complete genome sequence of the gastric pathogen Helicobacter pylori.
A:Reference number: A64520; MUID:97394467; PMID:9252185
A:Accession: F64543
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-502 <TOM>
A:Cross-references: UNIPROT:P56117; UNIPARC:UPI000013942F; GB:AE000539; GB:AE000511; NI;

C;Genetics:
A;Start codon: TTG
C;Superfamily: cardiolipin synthase

Query Match 23.1%; Score 627; DB 2; Length 502;
Best Local Similarity 31.5%; Pred. No. 3.2e-37;
Matches 156; Conservative 86; Mismatches 161; Indels 92; Gaps 15;

Qy	64	IRHTPHNNGLSDIY-----LLDDPHEALAAARAALIESAEHSLDLQYIWRN	109
Db	30	ISYDPYTPPGISLVAEKLKENPNHSAAILLEDGFALLHRVGLIRMSOKSIDMQYIYKN	89
Qy	110	DISGRLLFNLMYLAAGRVVRVLLDDNNTRGLD---DLILALDSHPNIEVRLNFPVL	165
Db	90	DLSSQVATAKELLNAARGVKVRIILLDDN---GLSDSDFSDIML-LNFHKNIEVKIENPYI	145
Qy	166	RKWRALGY---LTPFPRLNRRMHNKSFTADNRATILGGRNIGDEYFKVGEDTVFADLDIL	222
Db	146	RN-RGLRYFEMLADYERIKKEMHNKLFIVDNFAVIIGGRNIGDNYFDNDLDTNFDLDDAL	204
Qy	223	ATGSVGEVSHDFDRYVASHSAHNATRIIRSGNIGKGLQALGYNDETSRHALLRYETVE	282
Db	205	FFGVASKAKESFEYWRFHRSIPVS-LIRT-----HKELK-NNAKE	244
Qy	283	QSPUYQKIQTGRIDMQSVQTRL-----ISDTPAKGLDRDRRKPP	321
Db	245	IAKLHEKIPIISAEKNQKQPEKKNDFIDRFQKYQYPIYCGNAIFLADSPKK-IDTPLY-S-P	302
Qy	322	IAGRLQDALKOPEKSVLYLVSPYFVPTKSGTDALAKLVODGDVTVTLTNSLOATDVAVAHS	381
Db	303	IKIAFEKALQNAKQSVFIASSYFIPGKQWKIFKNQISKGIELNLTITNSLSTDAIVYG	362
Qy	382	GYVKYRKPPLKAGIKLYELOPNHVAVPATKKGLTGSSVTSLHAKTFIIVDGKRIFIGSNIL	441
Db	363	AWERYNQLVRMANGVNYEIRNDFNFRQIKGR---FSTKHSLGKGTIVFDNLTLLGGSFNI	419
Qy	442	DPRSARLNTENGVIIESPKIAEOWERTLADTTPYAVRYVTLDDKHNRLQWHDPAKTKYPN	501
Db	420	DFRSAYINTESAUVFDNFSFAKRV-----RUSLKDHAQQSWHLVVYR-----	461
Qy	502	EPEAKLMKRIAAKIL	516
Db	462	--HRVIEAVEEGIL	474

RESULT 12
B82724
cardiolipin synthase XF1087 [imported] - Xylella fastidiosa (strain 9a5c)
C:Species: Xylella fastidiosa
C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 16-Aug-2004
C:Accession: B82724
R:anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequencing
A:Nature 406, 151-157, 2000
A>Title: The genome sequence of the plant pathogen Xylella fastidiosa.
A:Reference number: A82515; PMID:10910347
A>Note: for a complete list of authors see reference number A59328 below
A:Accession: B82724
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-652 <SIM>
A:Cross-references: UNIPROT:Q9PEE1; UNIPARC:UPI000000C2600; GB:AEO03945; GB:AEO03849; NITD
A:Experimental source: strain 9a5c
R:Simpson, A.J.G.; Reinach, F.C.; Artuda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; A
Briones, M.R.S.; Bueno, M.P.; Camargo, A.A.; Canargo, L.E.A.; Carrato, D.M.; Carrier, H
as-Neto, E.; Docena, C.; El-Dorri, H.; Facincani, A.P.; Ferreira, A.J.S.

M.; Tshako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z.
A;Reference number: A59328
A;Contents: annotation
C;Genetics:
A;Gene: Xfr087
C;Superfamily: cardiolipin synthase

Query Match	16.1%	Score 439	DB 2	Length 652
Best Local Similarity	28.0%	Pred. No. 1.5e-23		
Matches 161	Conservative 69	Mismatches 203	Indels 144	Gaps 21
QY	47	SRHENTKPVLLDNLILQIRHPTPHNNGLSDIYLLDDPHEALAAARAAALIESAEHSILDLOYI	106	
DB	61	SRALSESTPV-----TP-----LHVVTILDKGESALVARINILRSATRIDMTYI	106	
QY	107	WRNDISGRLLFNLMYLAERGVRVRELLDDNNTRGLDILLAL-DSHPNIEVLFNPFVL	165	
DB	107	FDKDSARLIIMDELLTAARRGVRVRELLIDQLSAISDLNILGALAGAHVNFQRIYNPIFG	166	
QY	166	RKWRALGY-----LTDFFPLNRMRHKNKSF7ADNRATILGCGNIGDEYFKVGEDTVFAD	218	
DB	167	K-AKLANGYDYVASVLCCEFRFNQBMHKKLLVIDEMIGVVGGRNYQDDYVDWDLNFRD	224	
QY	219	LDILATGSGVGEVSHDFRYWASHSAHNATRIIRSONICK-----GLQALGYNDETSRHA	273	
DB	225	RDVLVAGPVLQMAVNFDAFWA--AERSVPVECLRDVGRMVLQDGVFML---PSAVFHP	278	
QY	274	LLR-YRETVE--OSPLYQK-----IQTGRIDWQSVQTRLISDTPAKGLDRDRRK-P	323	
DB	279	LERVQRTAEANDPDYVKRTFVDPALAVNKVQY-----VADLPQKRRHEHNANPVSIG	331	
QY	324	GLQDALQKPEKSVYLVSPYFPTKSGTDTALAKL--VODGIDVTYLTNSIQATDVAAVHS	381	
DB	332	PRLDLSINARHEVILQPIYLVLSKPALNIFRLNARTQDKPRVVVVTNSLAATDNPVYA	391	
QY	382	GYKY-RKPLLKAGIKLYELOP-----	402	
DB	392	LFYKYKRNMEDLGFDDIYEYKFPQPONPPIDLTGVVPMGWNDSLOTQDRLAAAKVSD	451	
QY	403	---NHAVPATKDKGLTGSSV-----TSLHAKTFIVDGK	432	
DB	452	SDNARQLPSENKGGQVDRVLRTETPRPPLGIHTVKNKPLPVTTPGARMGLHAKSLVVDRK	511	
QY	433	RIFIGSNLDRPSARLNTMGWVIESPKIAEQMERTL-ADTTPYAVRVTLDKHNRLOWH	491	
DB	512	VGVLGTHNFDPRGSYNTAEVVVIEDPAFARLLASSIEGMDPCGNAWV-----	560	
QY	492	DPATRKTYP-----NEPEAKLWKRIAANKILSLP	520	
DB	561	--APRKLLPGHYKLNYSVGKLSE--ALPVLDDWP	590	

RESULT 13

G70063

cardiolipin synthase homolog ywnE - Bacillus subtilis

C/Species: Bacillus subtilis

C/Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 09-Jul-2004

C/Accession: G70063

R/Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Bertero, C.; Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Cho, A.; Ehrlich, S.D.; Emmerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E. Nature 390, 249-256, 1997

A/Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Galler, Iech, J.; Harwood, C.R.; Henauf, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.P. Koether, P.; Koningstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardiniois, A.; Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maueel, Y. M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle, Rieger, M.; Rivolta, C.; Rochna, E.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanlon, A.; Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Sertor, akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpsira, P.; Tognoni, A.; Tosato, V.; Uchiyama, T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K. A.; Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A. A.; Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.

Qy	365	TVLTNSLOATDVA	VHSGVYKRP	LLKAGIKLYELQPNH	AVPA7KDKGLTGSSV	TLHA	424
Db	361	RVMIPN--KPDH	FVYRATTNYAGEL	METGAKIF-----	YDNGF-----	IHA	401
Qy	425	KTFIVDGKRI	FIGSNNLDP	RSARLNTENG	VIESPKIAEQMERT	IAD-----	TPPE-Y 476
Db	402	KTLVVDGEI	ASVGTANM	DFRSLNF	FEVNAFIERK	WVQKLEDAF	LEDILKSYQLTPELY 461
Qy	477	AYR	479				
Db	462	AKR	464				

RESULT 18
AG1387
cardiolipin synthase homolog lmo2503 [imported] - *Listeria monocytogenes* (strain EGD-e)
C/Species: *Listeria monocytogenes*
C/Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 09-Jul-2004
C/Accession: AG1387
R/Glasz, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecker, H.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.; D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A/Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurepkat, G.; Madueno, E.; Maitournam, A.; Makok, C.; Schluter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland, A./Title: Comparative genomics of *Listeria* species.
A/Reference number: AB1077; MUID:21537279; PMID:11679669
A/Accession: AG1387
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-482 <GUA>
A/Cross-references: UNIPROT:Q8Y4E3; UNIPARC:UPI0000055715; GB:NC_003210; PIDN:CAD00581.1
A/Experimental source: strain EGD-e
C/Genetics:
A/Gene: lmo2503
C/Superfamily: cardiolipin synthase

[illegible]

RESULT 19
G89906
hypothetical
C:Species
C:Date: 1
C:Accession
R:Kuroda,
ma, A.; M
C.; Shib
Lancet 35
A>Title: 1
A:Referen
A:Accessi
A>Status:
A:Molecu
A:Residue
A:Cross-r
A:Experi
C:Genetic
A:Gene. S
C:Superfa

Query Match 10.5%; Score 284.5; DB 2; Length 505;

Best Local Similarity	21.7%;	Pred. No. 1.3e-12;	
Matches	103;	Conservative	79;
Mismatches	158;	Indels	135;
Gaps	15		

QW 63 Q1P2T2D2NN/C1 S2T1V1 I1N2D2E2C1 A2D2A2 I2C2A2D2U2C1 D2 O2V2T2D2M2D2T2C2D2 I2 E2N2 I2M2V2 122

[illegible]

DB 143 RVDHFIDGNDLYDQVLXD-----IKNKEYIHLEYTYTFALDGLGRILHALEE 190

QY 123 AAERGVRLLDDNNTRGLDDLLALDSDHNIEVLFNPFVLRKWRALG-----YLTLD 176

Db 191 KLKQGLEVKILYDDVGSK-----NVKMANFDHF-----KSLGGEVEAFFASK 232

QY 177 FP----RLNRRMHKSFADNRATILGGRNIGDEYFKVGEDTVFADLDILATGSVVGVS 232

```

Db      233  LPLNFRMNRNRHRTIVIDGQLGYVGGFNIGDEYLGGLKGYWRDTHLFIQGDAVDALQ 292

```

Qv 233 HDEFYVASHSAHNATRII-----RSGNICKGLOALGYNDETSRHALI.RYRETYEOSP 285

[illegible]

DD URG 1000WS - QATKRFQZ1DVNR1FFNNNGFTU-----NSP 320

QY 286 L Y Q I Q T G R I D W Q S V Q --- T R L I S D T P A K G L D R D R R K P P I A G R L Q D A L K Q P E K S V Y L V S P 342

Db 327 IQIAAGPASPWHQIEYGYTKMMSA-----KKSVLQSP 361

Qy 343 YFVPTKSGTDALAKLVQDGDIDVTVLTNLSQATDVAAVHSGYVYKRPILKAGIKLYELQP 402

D**b** 362 YFIPD**NS**Y**IN**A**K**I**A**K**S**G**V**D**V**H**L**M**I**P**C**--K**P**D**H**P**L**V**Y**W**A**T**S**N**A**S**D**L**L**S**S**G**V**K**I**Y----- 415

QY 403 NHAVPATKDKGLTGSSVTSLSHAKTFIVDGKRIFIGSFNLDPPRSARLNTENGWIESPKIA 462

416 -----IHSKMCCTJDDIETVSGTANMDPSPSEI NEEFWAEIFYDENA 463

[illegible]

463 EQMERLADTTPETAYRVTLDKHNRLQWHDFAIRKTYTPNEPEAKLWRRFAAKILS 51/
QY :: || : | : | ::|| : : ||::||

Db 463 KDL-----RVAYHDITKSKQL-----TKESYANRPLSVKFESLAKLVS 502

RESULT 20
H9C74A

cardiolipin synthase [imported] - *Lactococcus lactis* subsp. *lactis* (strain IL14 n80/44)

C;Species: *Lactococcus lactis* subsp. *lactis*
C;Date: 23-Mar-2001 #sequence revision 23-Mar-2001 #text change 09-Jul-2004

C;Accession: H86744
R:Polotin, A : Wincker, P.: Mauger, S.: Jaillon O.: Malarme K.: Weissenbach

Genome Res. 11, 731-753, 2001

A; title: the complete genome sequence of the lactic acid bacterium *Lactococcus*
A; Reference number: A86625; UID:21235186; PMID:11337471

A;Accession: H86744
A;Status: preliminary

A:Molecule type: DNA	
A:Residues: 1-481 <STO>	
A:Cross-references: UNIPROT:Q9CGY2; UNIPARC:UPI000000C695C; GB:AE005176; PID:gl2723898; F	
A:Experimental source: strain IIL403	
C:Genetics:	
A:Gene: cIsA	
C:Superfamily: cardiolipin synthase	
Query Match	10.4%; Score 282.5; DB 2; Length 481;
Best Local Similarity	22.5%; Pred. No. 1.7e-12;
Matches 104; Conservative	76; Mismatches 167; Indels 115; Gaps 15;
QY	70 NNGLSDIYLLDDPHEALARAALIESAEHSLDLOQYIWRNDISGRLLFNFLMYLAAERQVR 129
DB	118 NTG----VKIFNDGRQKFNALIDDINKAKHHVMEYIIFRDMRGHEIYDALLAAKRGVE 174
QY	130 VRLLLD---DNNTRGLDLLLLALDSHPNIEVRLFPVLRKWRALGYLTDF----- 177
DB	175 VKLLIDANGSNKT-----MSNFKELO-----EAGGHVAQFFPLIMPLVN 214
QY	178 PRLNRMHNKSFADNRATILGGRNIGDEYFKVGED-TVPADLDILATGSGVVGEVSHDFD 236
DB	215 PRNTRYLRHKIIVIDGIIYTGGENGVGEYASITKKFGYWRDNHLRLTGDIVYLSQHRFI 274
QY	237 RYWASHSAHNATRIIRSGNIGKGLQALGYNDETSRHALLRYRETVEQSPLY-QKIQTGRI 295
DB	275 LDW-----NSQH-----HFEVTEGEPYPDSIVEGH- 300
QY	296 DMQSVQTRLIISDTPAKGLDRDRKPPFIAGRLQDALKQPEKSVYLVSPYFVPTKSGTDALA 355
DB	301 ---VATQLVTSGP-----DEDKEQIKLYMKMISGAEREIIQTPYVYIPSDALHESIK 350
QY	356 KLVDQGDIVTUNSLQATDVAAVHSGYVKYRKPLLKAGIKUYELOPNHAVPATKKGILT 415
DB	351 LALLSGVQVKLLIPN--RPDHLVYMATYFHAADLVKYGAKY-----TYENGFG- 397
QY	416 GSSVTSLHAKTIVDVKRIFGTGSFNLDPRSARLNTMGVWIESPKIAEOMERTIADTTPE 475
DB	398 -----VHSKTLIIIDGEVASGSANLDYRSLOLCEANVVIYDYDSOKLRN----- 443
QY	476 YAYRVTLDKHNRLOWHDPATRKTYNEPEAKLWKRIAAKILS 517
DB	444 -----DFMKDLKLSRPLTLERYEERSKLVRFKEGLARLIA 478
RESULT 21	
B84007	
cardiolipin synthetase BH2858 [imported] - Bacillus halodurans (strain C-125)	
C:Species: Bacillus halodurans	
C>Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 09-Jul-2004	
C:Accession: B84007	
R;Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fuji, F.; Hira	
Nucleic Acids Res. 28, 4317-4331, 2000	
A:Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and	
A:Reference number: A83650; MUID:20512582; PMID:11058132	
A:Accession: B84007	
A>Status: preliminary	
A:Molecule type: DNA	
A:Residues: 1-503 <STO>	
A:Cross-references: UNIPROT:Q9K824; UNIPARC:UPI0000127B9A; GB:AP001516; GB:BA000004; NID	
A:Experimental source: strain C-125	
C:Genetics:	
A:Gene: BH2858	
C:Superfamily: cardiolipin synthase	
Query Match	10.1%; Score 276; DB 2; Length 503;
Best Local Similarity	24.5%; Pred. No. 5.3e-12;
Matches 110; Conservative	57; Mismatches 188; Indels 94; Gaps 12;
QY	61 LIQIRHTPHNGLS---DIYLLDDPHEALARAALIESAEHSLDLOQYIWRNDISGRLLF 117
DB	124 LFRLAHLRANNPISFTNTKVLTDGKETFAHIKQALRMATHIHLEXYIVRDEIGQBIK 183

QY	118 NLMYLAAERGVRVRLLLDDNNTRGLDLLLLALDSHPNIEVRLFPVLRKWRALGYLTDF 177
DB	184 EILMOKAKEGIHVRFLVDGVGSWKLSKSYIQDLKQAGVEIVFPAPVKL----- 231
QY	178 P-----RLNRRMHKSFADNRATILGGRNIGDEYFKVGEDTVFA---DLDLATGSGVVGE 230
DB	232 PFINHTINYRNHRKIIIVIDGTGVFGVGLNIGDEY--LGKDPYFGFWRDTHLYVRGEAVRT 289
QY	231 VSHDFDRYWASHSAHNATRIIRSGNIGKGLQALGYNDETSRHALLRYRETVEQSPLYQKI 290
DB	290 LQLIFLRDWAHETG-----ETILKPSYLSPA 315
QY	291 QTGRIDMQSVQTRLIISDTPAKGLDRDRKPPFIAGRLQDALKQPEKSVYLVSPYFVPTKSG 350
DB	316 LTNMKDDGGVQ--MIASGP-----DYRWEINKKLPFSMITSAKKSIWITSYPFIPDEDI 367
QY	351 TDALAKLVQDGDIVTUNSLQATDVAAVHSGYVKYRKPLLKAGIKUYELOPNHAVPATK 410
DB	368 LSALKIAALSGIDVRILVPN--RPDKRIVFHASRSYFPELLEAGVKVVEY----- 415
QY	411 DXGLTGSVTSLHAKTIVDVKRIFGTGSFNLDPRSARLNTMGVWIESPKIAEOMERTLA 470
DB	416 TRGP-----LHSEKIIIVDNEIASIGTSNMDMRSFHLNFEVNAF-----LYRTKS 459
QY	471 DTTPEYAVRVTLDKHNRLOWHDPATRKTY 499
DB	460 VTTLVSDFVDVLEHTNQIRFEQFNRAWY 488
RESULT 22	
C84125	
cardiolipin synthetase BH3803 [imported] - Bacillus halodurans (strain C-125)	
C:Species: Bacillus halodurans	
C>Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 09-Jul-2004	
C:Accession: C84125	
R;Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fuji, F.; Hira	
Nucleic Acids Res. 28, 4317-4331, 2000	
A:Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and	
A:Reference number: A83650; MUID:20512582; PMID:11058132	
A:Accession: C84125	
A>Status: preliminary	
A:Molecule type: DNA	
A:Residues: 1-398 <STO>	
A:Cross-references: UNIPROT:Q9K6C6; UNIPARC:UPI000000C42DC; GB:AP001520; GB:BA000004; NID	
A:Experimental source: strain C-125	
C:Genetics:	
A:Gene: BH3803	
C:Superfamily: cardiolipin synthase	
Query Match	10.1%; Score 275; DB 2; Length 398;
Best Local Similarity	23.9%; Pred. No. 4.4e-12;
Matches 108; Conservative	63; Mismatches 172; Indels 108; Gaps 14;
QY	65 RHTPHNGLSDIYLLDDPHEALARAALIESAEHSLDLOQYIWRNDISGRLLFNLMYLA 124
DB	35 RNQPIRHG--IVRLLPDTCDEFIGSLIHDIAKAAEHIIHILFYIFRDDHIGKILQHEQKA 92
QY	125 ERGVVRVRLLD-----DNNTRGLDLLLLALDSHPNIEVRLFPVLRKWRALGYLTDFPRL 180
DB	93 KEGVAVRLLVDRFGADVKNRSIQSLKQA-GAOPYAHRISFPY----W-----PFSL 139
QY	181 NRRMHKSFADNRATILGGRNIGDEYFKVGEDT---VFADLDILATGSGVVGEVSHDFDR 237
DB	140 NRRNHRKITVDGKIYIGGVNIGDEY--LGRDPKLGFWRDYHLRLTGDGVQDLODQFIQ 197
QY	238 YWASHSAHNATRIIRSGNIGKGLQALGYNDETSRHALLRYRETVEQSPLYQKIQTGRIDW 297
DB	198 DWERESRLPVTR----- 209
QY	298 QSVQTRLIISDTPAKGLDRDRKPPFIAGRLQDA----LKQPEKSVYLVSPYFVPTKSGTDA 353
DB	210 ----DRWLYPPPLAKGPHELRIIPTNGSFLSDLSQLVEQAEETITIGTFYFIPGKLLHA 265

QY 354 LAKLVQDGDVTLVNSLOATDVAHVSGYVYKRPKLLKAGIKLYELOPNHAVPATKDKG 413
Db LLDAAARGVAVRLVLP--KKGDPVLVKEAAPFYFKELLEGGINIYRY-------RG 313
QY 414 LTGSSVTSLHAKTFTVDGKRIFIGSNLDRPSARLNTMGVVIESPKIAEQMERTLADTT 473
Db 314 F-----FHGAIVDDKLADVGTANFDKRSFRLNVEINCLLYDKEMIQVLVBEL----- 362
QY 474 PEYAYRVTLDKHNLQWHDPPATRTKTYNEPE 504
Db 363 -DYDFSIS---ERLQMEDLANRSPFFHRTKE 388

RESULT 23
T43863
cardiolipin synthase [imported] - Clostridium perfringens
C:Species: Clostridium perfringens
C:Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 09-Jul-2004
C:Accession: T43863
R:Koyama, M.
submitted to the EMBL Data Library, August 1998
A:Description: Clostridium perfringens hem operon.
A:Reference number: 222707
A:Accession: T43863
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-476 <KOY>
A:Cross-references: UNIPROT:Q9ZNC6; UNIPARC:UPI000016EODF; EMBL:AB017186; PIDN:BAA74786.
A:Experimental source: strain NCTC8237; vegetative cell
C:Genetics:
A:Gene: cls
C:Superfamily: cardiolipin synthase

Query Match 10.0%; Score 273; DB 2; Length 476;
Best Local Similarity 24.3%; Pred. No. 8e-12;
Matches 108; Conservative 79; Mismatches 167; Indels 90; Gaps 20;

QY 41 LEERTSRHNTSKP-VLLDN-----LQIRHTPHNGL-----SDIYLLDDPHEALAA 89
Db LDDTKQKYLNSFKSHYKLDNISLKYKDLIMNPNNDNSTVTQRNDIDLFDANSLFEEM 125
QY 90 AALIESAEHSLDLOYIWRNDISGRLLFNLMYLAERGVRVRLLLDD--NNTRGLD-DLL 146
Db 126 IDEINKAEKTIHMEFYIFKSDIEIGKKILOALTCKAKEGVEVKLVDSIGNSIHKKOIDKL 185
QY 147 LALDHPNIEVRLNPFVLRKWRALGYLTDFPRLNRHMHKSFADNRATILGSGNIGDE 206
Db 186 KAAGG---DFKIFPGFCK-----YIN--LRINYRNHRKILIDSKVAFLOGFNGID 233
QY 207 YFKVGEDTVFADLDILATGVSVGVSHDFORYWASHSAHNATRIIRSGNI--GKGLQALGY 265
Db 234 Y--LGDK-----NIGH-----W--RTHYKIKGLAINDLEGRLDWSY 269
QY 266 NDETSRIA--LLRYRETVEQSPLYOKIQTGRIDWOSVOTRLISDTPAKGLDRDRKPPPIAG 324
Db 270 ANESDLIDLKXYFINPSTDLPKI-----IGAQIVSSGP-----DHTEQOIKN 314
QY 325 RLQDALQPEKSVYLVSPYFPTKSGTDLAKLVODGIDVTVLNLSQATDVAHVSGYV 384
Db 315 GYFKIINSKKNFIQTPYFVDPPEMLEALRLAALSQVDVKIMLPF---NPDHKFMGWI 370
QY 385 --KYRKPLKAGIKLYELOPNHAVPATKDKGLTGSSTLSHAKTFTVDGKRIFIGSNLD 442
Db 371 ANSYFESLLNAGAKIY-----LYEKG-----LHAKTIVADSSICSVGTANMD 413
QY 443 PRSARLNTMGVVIESPKIAEQME 466
Db 414 IRGFSLNFESNIFIYNEAISKSME 437

RESULT 24
H82711
cardiolipin synthase XFI209 [imported] - Xylella fastidiosa (strain 9a5c)

C:Species: Xylella fastidiosa
C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004
C:Accession: H82711
R:anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequencing
Nature 406, 151-157, 2000
A:Title: The genome sequence of the plant pathogen Xylella fastidiosa.
A:Reference number: A82515; MUID:20365717; PMID:10910347
A:Note: for a complete list of authors see reference number A59328 below
A:Accession: H82711
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-467 <SIM>
A:Cross-references: UNIPROT:Q9PE19; UNIPARC:UPI00000C265D; GB:AE003954; GB:AE003849; NI
R:Experimental source: strain 9a5c
R:Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.;
Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrez,
as-Neto, E.; Docena, C.; El-Dorri, H.; Facincani, A.P.; Ferreira, A.J.S.
submitted to GenBank, June 2000
A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Froh
J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laig
chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins,
A:Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.
F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.
Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasa
A:Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silve
M.; Tshako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.;
A:Reference number: A59328
A:Contents: annotation
C:Genetics:
A:Gene: XFI209
C:Superfamily: cardiolipin synthase

Query Match 9.9%; Score 270.5; DB 2; Length 467;
Best Local Similarity 27.3%; Pred. No. 1.2e-11;
Matches 117; Conservative 56; Mismatches 162; Indels 93; Gaps 18;

QY 78 LLDPDHEALAAALIESAEHSLDLOYIWRNDISGRLLFNLMYLAERGVRVRLLLDD- 136
Db 114 LLDYQSSRLLLNDIDSACERVHLLYLMFDDAVGDAVVAALQRAAARGVSCRLLDVAV 173
QY 137 NNTRGLDLALLDHPNIEVRLNPFVLRKWRALGYLTDFPRLNRHMHKSFADNRAT 196
Db 174 GAKRGLRAYQCCILRAH-GVEVHAMLPGLR-WRRSG-----RMDLRNHRKIAVIDNRVA 225
QY 197 ILGRNIGDEYFKVGEDTVFADLDILATGVSVG-EVSHDFDRYWASHSAHNATRIIRSGN 255
Db 226 YIGSQNLACPDFV-----PHFPNRELVA--RVCQPAVSH----- 257
QY 256 IKGLOALGYND---ETSRHALLRYRETVEQSPLYOKIQTGRIDWOSVOTRLISDTPAKG 312
Db 258 ----LEAVFLSDWFVETGQR--LRMSDV---PICE-----QNVAAQLVASGPAYP 299
QY 313 LDRDRKPPITAGRLQDALQPEKSVYLVSPYFPTKSGTDLAKLVODGIDV--TVLTNSL 371
Db 300 YENAR-----DAVNTMIHLARRRVLTTPYFVDPDEATLSALRIAGTSGVDVQLILSEN 353
QY 372 QATDVAHVSGYVYKRPKLLKAGIKLYELOPNHAVPATKDKGLTGSSTLSHAKTFTVDG 431
Db 354 NOFLMAWAQEAFF---EELLRAGIKALYRPHF-----LHAKHLSVDD 393
QY 432 KRIFIGSNLDRPSARLNTMGVVIESPKIAEQMERTLADTTPEYAYRVTLDKHNLQW 491
Db 394 DIALFGSINLDRSFALNAEIGMICYDSRIVSLREVESD-----YLANACQLNLVEWQ 447
QY 492 D-PATRT 498
Db 448 QRPWERS 455

RESULT 25
B82971
cardiolipin synthase PA5394 [imported] - Pseudomonas aeruginosa (strain PAO1)
C:Species: Pseudomonas aeruginosa

C>Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2004
C/Accession: B82971
R/Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; B
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,
.; Lory, S.; Olson, M.V.
Nature 406, 953-964, 2000
A>Title: Complete genome sequence of *Pseudomonas aeruginosa* PA01, an opportunistic patho
A/Reference number: A82950; MUID:20437337; PMID:10984043
A/Accession: B82971
A>Status: preliminary
A/Molecule type: DNA
A/Residues: 1-490 <STO>
A/Cross-references: UNIPROT:Q9HTH0; UNIPARC:UPI00000C5PFF; GB:AE004952; GB:AE004091; NID
A/Experimental source: strain PA01
C/Genetics: cis; PA5394
C/Superfamily: cardiolipin synthase

Query Match 9.9%; Score 268; DB 2; Length 490;
Best Local Similarity 24.5%; Pred. No. 1.9e-11;
Matches 111; Conservative 60; Mismatches 186; Indels 96; Gaps 13;

Qy 37 WLPLEERTEGSHFNTSPVLNDNIQIRHTP--HNGLSDIYLLDDPHEALAAALIE 94
Db WRPVEEALAAQVSGYKG--LKALVRMTRTPTLANN---RVRLLVNGEASFAMFKAIS 151
Qy 95 SAEHSLDLQYIWRNDISGRLLFNLMYLAARGVVRLLDDNNTRGLDILLALDSHPN 154
Db AARQVILVQFFIVRDDALGQRLQQLLERAANGVEVFFLYDAIGSHALPHRYVERLRQGG 211
Qy 155 IEVLFNFPVLKRWALGYLTFPRLNRRMNKSPADNRATILGGRNIGDEYFKVGEDT 214
Db VQMHGFS-----TGSGLNRRF-QVNFNHRKVVVVVDEGCFVGHNVGVEY--LGKRP 261
Qy 215 VFA---DLIDLATGSGVGEVSHDF--DRYWAHSAHNATRIIRSGNIGKGLQALYNDET 269
Db PLAPWRDTHMELRGPAVCLQBSFAEDWYATHSL----- 296
Qy 270 SRHALLRYRETVESQPLVQKIQOTGRIDWQSVQTRILISTPAKGLDRDRKPPKIAGRLQDA 329
Db PPLILPPQVDSGALCQVVASGPAQAQETCSLF-----FVEM 333
Qy 330 LKQPEKSVVLVSPYVPVTKSGTDALAKLVQDGIDVTVLNTSLQATDVAAVHSGYVKRKP 389
Db INAAHERVWITSFYVPDEAVMAURLAVLRGVDRLLIPS--RPDHTVTVAASLYALE 391
Qy 390 LKAGIKLYELOPNHAVPATKDKGLTSSVTSLSHAKTFIVDGKRIFIGSFNLDPRSARLN 449
Db AIRAGVKVFXYQPGF-----LHQKVVLDVDRDTAAVGSANLDNRSFRLN 434
Qy 450 TEMGVVIESPKIAEQWERTL-----ADTTPE 475
Db FEVMVTVDEGPAGEVAMLEADFAESLEFTE 467

RESULT 26
S60089
cardiolipin synthetase homolog ywiE - Bacillus subtilis
C/Species: Bacillus subtilis
C/Date: 15-Feb-1996 #sequence_revision 01-Mar-1996 #text_change 09-Jul-2004
C/Accession: S60089; S55414; F70059
R/Cruz Ramos, H.; Boursier, L.; Moszer, I.; Kunst, F.; Danchin, A.; Glaser, P.
EMBO J. 14, 5984-5994, 1995
A>Title: Anaerobic transcription activation in *Bacillus subtilis*: identification of disc
A/Reference number: S60080; MUID:96112813; PMID:8846791
A/Accession: S60089
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A/Molecule type: DNA
A/Residues: 1-500 <CRU>
A/Cross-references: UNIPROT:P45860; UNIPARC:UPI0000060BCF; EMBL:Z49884; NID:9971335; PID
R/Glaser, P.; Danchin, A.
submitted to the EMBL Data Library, May 1995

A/Description: Cloning and sequencing of the *Bacillus subtilis* chromosomal region from 3
A/Reference number: S55414
A/Accession: S55414
A>Status: preliminary
A/Molecule type: DNA
A/Residues: 432-500 <GLA>
A/Cross-references: UNIPARC:UPI000016B83F; EMBL:Z49782; NID:9853752; PIDN:CAA89861.1; PFI
R/Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berter
C.; Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Cho
A.; Ehrlich, S.D.; Emmerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.
Nature 390, 249-256, 1997
A/Authors: Foulger, D.R.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Galler
iech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.
Koetter, P.; Konigstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois
Y. M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Levine, A.; Liu, H.; Masuda, S.; Maueel
Rieger, M.; Rivolta, C.; Roche, E.; Roche, M.; Rose, M.; Sadale, Y.; Pohl, T.M.; Portetelle
A/Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Seror
akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama,
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K.
A/Authors: Yoshikawa, H.F.; Zumbstein, E.; Yoshikawa, H.; Danchin, A. *Bacillus subtilis*.
A>Title: The complete genome sequence of the Gram-positive bacterium *Bacillus subtilis*.
A/Reference number: A69580; MUID:98044033; PMID:9384377
A/Accession: F70059
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A/Molecule type: DNA
A/Residues: 1-500 <KUN>
A/Cross-references: UNIPARC:UPI0000060BCF; GB:Z99123; GB:AL009126; NID:g2636240; PIDN:CA
A/Experimental source: strain 168
C/Genetics: ywiE
C/Superfamily: cardiolipin synthase

Query Match 9.8%; Score 266.5; DB 2; Length 500;
Best Local Similarity 22.4%; Pred. No. 2.5e-11;
Matches 107; Conservative 64; Mismatches 176; Indels 131; Gaps 15;

Qy 48 RHFNTPKVLNDNIQIRHTPHNGLSD-----TYLLDDPHEA 85
Db RMYNREK---URKULDKETEPTVTLKDNQERFTYSIRAAHMINTKSNIVKLANGET 150
Qy 86 LAARAAALIESAEHSLDLQYIWRNDISGRLLFNLMYLAARGVVRLLDDNNTRGLDDL 145
Db PPDIFPKAMRKAESYIHIEYMFKSDMLGRGMDIMMEKARQGEVRFYDAAGSMKLARR 210
Qy 146 LLALDSHPNIEVLFNFPVLKRWALGYLTFPRLNRRMNKSPADNRATILGGRNIGD 205
Db DIMRMKQAGVDIVPFSPL-----KYGFFNQ--KLNFRNHRKVIIDGKTGFVGLNVGK 262
Qy 206 EYFKVGEDTVPADLLDILATGSGVGEVSHD-PDRYWAHSAHNATRIIRSGNIGKGLQALG 264
Db EY-----ISRDPYIGWRD-----THLRLEGIVQTLHAIF 293
Qy 265 YND--ETSRHALLRYRE-----TVEQSPLYQKIQOTGRIDWQSVQTRILISTPAKGLDRDR 317
Db MLDWEYVSNVILDOEEYNTVPVVEGGIYQIVATG-----PD 331
Qy 318 RKPPIAGRLQDALQPEKSVVLVSPYVPVTKSGTDALAKLVQDGIDVTVLNTSLQATDVA 377
Db MKESMSDLYEMISSAQKSIWIATPFYFPNIESIRTKAAATKGVVEVWVMP--EKNDSF 389
Qy 378 AVHSGYVKYRPLKAGIKLYELOPNHAVPATKDKGLTSSVTSLSHAKTFIVDGKRIFIG 437
Db LTQVASSYFPPELLLEGIEVYSYQ-----KGF-----MHQKVMIDGDLASVG 432
Qy 438 SFNLDPRSARLNTEMGVVI-----ESPKIAEQ--MERTLADTPE 475
Db TANMDRSFQLNFEVNVFFTDAAIRLTLEAHFEEDMQESELSPVGFYKRGVADRTEK 490

RESULT 27
AE3539
cardiolipin synthetase (EC 2.7.8.-) [imported] - *Brucella melitensis* (strain 16M)

C;Species: Brucella melitensis
C;Date: 01-Feb-2000 #sequence_revision 01-Feb-2002 #text_change 09-Jul-2004
C;Accession: AE3539
R;DelVecchio, V.G.; Kapatral, V.; Redkar, R.J.; Patra, G.; Mijer, C.; Los, T.; Ivanova,
.; Mazur, M.; Goltzman, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Letess
Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002
A;Title: The genome sequence of the facultative intracellular pathogen Brucella melitensis
A;Reference number: AD3252; PMID:1175668
A;Accession: AE3539
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-492 <KUR>
A;Cross-references: UNIPROT:Q8YDD8; UNIPARC:UPI00000583DD; GB:AE008918; PIDN:AAU53480.1;
A;Experimental source: strain 16M
C;Genetics:
A;Gene: BMEI0239
A;Map position: 11
C;Superfamily: cardiolipin synthase
C;Keywords: transferase

Query Match 9.7%; Score 263.5; DB 2; Length 492;
Best Local Similarity 23.0%; Pred. No. 4.1e-11;
Matches 98; Conservative 69; Mismatches 180; Indels 79; Gaps 13;

QY 48 RHNTSPVLLDNLQIRHPTHNNGLSDIYLLDPHEALARAALIESAEHSLDLQYIIV 107
DB 106 RQFGAMK-ILGDVSLYDFTSGNT-----IEMLEAGDEAYAAMLGAIGRAERSIVLEYIIF 160
QY 108 RNDISGRLLFNMLYLAERGVRLVLLDDNNTRGLDLDLALDHPNIEVLENFVLRK 167
DB 161 DHDIAIGKFFADALGDVQGVREVLVDVAGARYSPFSIVKLLKEKGVKAVFNGNIIG 220
QY 168 WRALGYLTDPRLNRRMHNKSFADNRATILGGRNIGDEYFK-VGEDTFEADLDILATGS 226
DB 221 LR-----LPEANLRTHRWMLIVDGTAFGGNNIRAGFVRAIAGDAVAFDFHKEGP 273
QY 227 VGVESHDFRYWASHAHNATRIIRSGNTCKGLQALGYNDTSRHALLRYRETVEOSPL 286
DB 274 AIADLFHIAEDMR-----FATGELLTG-----EAWSTAPP 304
QY 287 YQKIQGTGRIDWQSVOTRLSDTPAKGLDRDRKPPPIAGRLQDALKQPEKSVLYSPVFP 346
DB 305 ENPPGTGTL-----VAVWGGPDKNLETNHR-----MMGAFSIAQQHILIMTPYLLP 352
QY 347 TKSQTDALAKLVQDGDIVTL---TNSLQATDVAHVHSGVYKRPKLLKAGIKLYELQPN 403
DB 353 DRELISALVTAARGVSDIVVPGVNNKLVD-RAMPARF-----DQLLRDGCRIWR---- 403
QY 404 HAVPATKDKGLTGSSVTSIHAKTFFIVDGKRFIFGSFNLDPRLSRARLNTMGVIESPKIAE 463
DB 404 -----AGGAFFN--HSLKMTIDGMSYVGVSSNIDPRSLRLNFEVLEILDRDVAR 450
QY 464 QMERTL 469
DB 451 QVEERI 456

RESULT 28
D83103
probable phospholipase PA4339 [imported] - Pseudomonas aeruginosa (strain PA01)
C;Species: Pseudomonas aeruginosa
C;Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2004
C;Accession: D83103
R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; B
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,
.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A;Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic path
A;Reference number: A82950; MUID:20437337; PMID:10984043
A;Accession: D83103
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-359 <STO>

A;Cross-references: UNIPROT:Q9HW62; UNIPARC:UPI000000CSCCA; GB:AE004850; GB:AE004091; NI
A;Experimental source: strain PA01
C;Genetics:
A;Gene: PA4339
C;Superfamily: cardiolipin synthase

Query Match 9.6%; Score 261; DB 2; Length 359;
Best Local Similarity 23.8%; Pred. No. 3.8e-11;
Matches 93; Conservative 57; Mismatches 148; Indels 92; Gaps 10;

QY 93 IEAEHSLDLQYIWRNDISGRLLFNMLYLAERGVRLVLLDDNNTRGLDLDLALDLSH 152
DB 5 IEAQRSIELELYVEDHGHCAELFLVALDDARRRGVAVRCLFDGFGCLGLGSAWIOQLRE 64
QY 153 PNIEVLENFVLRKWRALG---YLTDFFRLNRRMHNKSFADNRATILGGRNIGDEYFK 209
DB 65 AGGELALYNPL---RWKLTGGNLY-----RDHKLILLVDGRLGVGGAGITDFEWE 112
QY 210 -VGEDTVFADLDILATGSSVVGVEVSHDFDRYWASHAHNA-----TRIIRSGNIGKG 259
DB 113 PVSIVSAWREVMVMDGPPVADWAALFEROMLACLEKAWKPREGMTLRLPPQGAARG 172
QY 260 LQALGYNDTSRHALLRYRETVEOSPLYQIKIQTGRIDWQSVOTRLSDTPAKGLDRDRK 319
DB 173 LGRVAYAD-----ARQHRDILQSLVRALNGSRRR 201
QY 320 PPIAGRLQDALKQPEKSVLYSPVFPVTKSGTDALAKLVODGIDVTVLTNSLOATDVAAV 379
DB 202 -----IWLATYFLEPTWKVRALRKAAGRGVEVRLLAG-RLTDHAPV 243
QY 380 HSGYVYKRPKLLKAGIKLYELQPNHAVPATKDKGLTGSSVTSIHAKTFFIVDGKRFIFGSF 439
DB 244 RYAGQRYPRLLRAGVRIHEYQPRF-----LHLKMMVVD-DWVSVGC 285
QY 440 NLDPRARLNTMGVIESPKIAEQMERTL 469
DB 286 NFDHNLRLFNLDANLEALDPDFTNEAASL 315

RESULT 29
E97307
probable cardiolipin synthase (phospholipase D family) [imported] - Clostridium acetob
C;Species: Clostridium acetobutylicum
C;Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 09-Jul-2004
C;Accession: E97307
R;Nolling, J.; Breton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee
.; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.
J. Bacteriol. 183, 4823-4838, 2001
A;Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Cl
A;Reference number: A96900; MUID:21359325; PMID:21359325
A;Accession: E97307
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-510 <KUR>
A;Cross-references: UNIPROT:Q97E04; UNIPARC:UPI00000D756C; GB:AE001437; PIDN:AAK81248.1
A;Experimental source: Clostridium acetobutylicum ATCC824
C;Genetics:
A;Gene: CAC3316
C;Superfamily: cardiolipin synthase

Query Match 9.6%; Score 260; DB 2; Length 510;
Best Local Similarity 20.8%; Pred. No. 7.6e-11;
Matches 103; Conservative 79; Mismatches 177; Indels 136; Gaps 16;

QY 36 SWLPPLERTESRHFTSKPVLNDNLQIRHPTHNNGLSDIYLLDDPHEALARAALIES 95
DB 138 SLPPPY-KYTKAEYFSTGETFNEELQLK-----X 167
QY 96 AEHSLDLQYIWRNDISGRLLFNMLYLAERGVRLVLLDDNNTRGLDLDLALDHPN- 154
DB 168 AKQYIFLEYFITKEGVMMNSVLQILDRKVOEGVEVVIYDD-----IGCMFTLPG 218
QY 155 -----IEVLENFVLRKWRALGYLTDPRLNRRMHNKSFADNRATILGGRNIGD 205

```
Db 219 YHKELEKIGKCCVFNPLI-----PWSF-KFNDRHRKIAVIDGLVGTGINLSD 269
Qy 206 EYF-KVGEDTVFADLDILATGSSVGEVGHDFDRYVASHSAHNATRIIRSGNIGKGLQALG 264
Db 270 EVINKYQYGVKWDFAIKVEGKAANLVSFMSWS-----FLR-GIDEDFNTFK 318
Qy 265 YNDETSRHALLRYRETVEQSPLYQKIQGRIDWQSVQTRLLSDTPAKGLDRDRKPPPIAG 324
Db 319 KNIKEFSRSREGYVQPFADSP-----DGEVPGEIYNMLIS----- 355
Qy 325 RLQDALKOPEKSVYLVSPYFPTKSGTDALAKLVQDGDVTLNLSQATDVAAVHSGYV 384
Db 356 -----KATKVYITTPYLVIGNEMVTALSAAGGVDRRIIPIH-----PKKIVHSTK 406
Qy 385 KYRPELLKAGIKLYELOPNHAVPATKDKGLTSSVTSLSHAKTFIVDGKRIFIGSFNLDPR 444
Db 407 SYKVLIESGVKIYEMPGF-----IHSKTYVCDNEYGVGVSINMDFR 449
Qy 445 SARLNTMGVWIESPKIAEQMERTLADTTPEYAYRVTLTKHNLQWHPDPAKTKYFNEPE 504
Db 450 SLYLHFECGVWMYKNTTYDIKDFMDTL-DKSKEITLERINKVKWY----- 495
Qy 505 AKLWK--RIAAIL 516
Db 496 STLWRVLRVAPLM 510

RESULT 30
AH2333
cardiolipin synthase [imported] - Nostoc sp. (strain PCC 7120)
C:Species: Nostoc sp. PCC 7120
A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Jul-2004
C:Accession: AH2333
R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi,
Nakazaki, N.; Shimo, S.; Sugimoto, M.; Takazawa, M.; Yanada, M.; Tabata, S
DNA Res. 8, 205-213, 2001
A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ana
A:Reference number: AB1807; MUID:21595285; PMID:11759840
A:Accession: AH2333
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-480 <KUR>
A:Cross-references: UNIPROT:Q8YPH1; UNIPARC:UPI00000CEAE63; GB:BA000019; PIDN:BAB75922.1;
A:Experimental source: strain PCC 7120
C:Genetics:
A:Gene: alr4223
C:Superfamily: cardiolipin synthase

Query Match 9.4%; Score 256; DB 2; Length 480;
Best Local Similarity 24.3%; Pred. No. 1.4e-10;
Matches 120; Conservative 69; Mismatches 181; Indels 124; Gaps 18;

Qy 35 SSWLPPLERTSRHFNTPKPVLLDNIILQIRHTPHNGLSDIYLLDDPHEALAAALIE 94
Db 98 SAKUEPLQ-----LLAEAFIGFTSGNNA-----KLLINGQQTVAWLSAIA 140
Qy 95 SAEHSLDIQYIWRNDISGRLLFNMLYLAERGVVRLLDDNNTGRGLDLLLLDLSHPN 154
Db 141 SANSYILLOSIVVDDKAGNEFKDALIAKAGIRVYLIYDEIGSNKISRLYVKSLOKYD 200
Qy 155 IEVRLFNFPVLRKWRALGYLDFPPLNRRMKNKSPADNRATILGGRNIGDEYFKVGBD- 213
Db 201 IQVSFAFHTTRGRGNRF-----QLNFRNHRKILVDGRTAFIGGLNISDEY--LGKNP 250
Qy 214 --TVPADLDILATGSGVGEVSHDF--DRYVASHSAHNATRIIRSGNIGKGLQALGYNDET 269
Db 251 RLSPWRDTHMLEGPTVOSLOGCFLODWYAT----- 282
Qy 270 SRHALLRYRETVEQSPLYQKIQGTGRIDWQSVQTRLLSDT-PAKGLDRDRKPPPIAGRL-- 326
Db 283 -----RQVIDVNWQVP-----NWESDYTALVFTTGA-----DKLK--ACKLFF 320
```

```
Qy 327 QDALKOPEKSVYLVSPYFPTKSGTDALAKLVQDGDVTLNLSQATDVAAVHSGYVKY 386
Db 321 VSAINQOQTRLMWATPYFEVDDSTLTALKLAALRGVDVRIILPN--RPDHLVLVLCFSFSY 378
Qy 387 RKPLLKAGIKLYELOPNHAVPATKDKGLTSSVTSLSHAKTFIVDGKRIFIGSFNLDPRS 446
Db 379 YTEKATNINKLYRYK--HGF-----MHQKVLIDKEWAGVGTVDLDRSF 421
Qy 447 RLNTE-MGVWIESPKIAEQMERTLADTTPEYAYRVTLTKHNLQWHPDPAKTKYFNEPE 505
Db 422 FLNFEVNGFVANSQFVKSEVLEQLADL--KAALAVDFSDYER-----K 462
Qy 506 KLVKRIAAIL 519
Db 463 YLWFKLAVRISLL 476

RESULT 31
C82171
cardiolipin synthase VCI670 [imported] - Vibrio cholerae (strain N16961 serogroup O1)
C:Species: Vibrio cholerae
C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004
C:Accession: C82171
R:Heidelber, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.;
Chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, P.
1, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
Nature 406, 477-483, 2000
A:Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
A:Reference number: AB2035; MUID:20406833; PMID:10952301
A:Accession: C82171
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-484 <HRI>
A:Cross-references: UNIPROT:Q9KRH2; UNIPARC:UPI00000C307E; GB:AE003852; NID
A:Experimental source: serogroup O1; strain N16961; biotype El Tor
C:Genetics:
A:Gene: VCI670
A:Map position: 1
C:Superfamily: cardiolipin synthase

Query Match 9.0%; Score 243.5; DB 2; Length 484;
Best Local Similarity 25.0%; Pred. No. 1.1e-09;
Matches 104; Conservative 59; Mismatches 142; Indels 111; Gaps 18;

Qy 71 NGLSDIYLLDDPHEALAAALIESAHSGLDQYIWRNDISGRLLFNMLYLAERGVV 130
Db 119 NTLG---LLSNPNEILHAIIDIERAQFQIRMVFIWHPGGLADAVASAVIQASKRGVNV 175
Qy 131 RLILDDNNT-----RGLDLLLLDLSHPNIEVRLFPVLRKWRALGYLTD 176
Db 176 KLLDLSAGSPRFFSPPEKMMRDAGI-EVQALEVSP-----WRIF----- 215
Qy 177 FPLNRNMHNSFTADNRATILGGRNIGD-EYFK---VGEDTVFADLDILATGSSVGEV 231
Db 216 LRRDLQHRKIIIVDDIEIATGSMNVDPAYFQKAGVQ--WIDIMVVTGTNNVL 272
Qy 232 S--HDFPRYVASHSAHNATRIIRSGNIGKGLQALGYNDETSRHALLRYRETVEQSPLYQK 289
Db 273 SAHCWD--MEFET-----GSRMLPKNPE-----CRLEPNQPHPI-QV 308
Qy 290 IOTGRIDWQSVQTRLLSDTPAKGLDRDRKPPPIAGRLDALKOPEKSVYLVSPYFVFKS 349
Db 309 VPSG---PGMPENLISQV-----LTLAINQANRSVCITTPYFVPSAD 347
Qy 350 GTDALAKLVQDGDVTLNLSQATDVAAVHSGVYKVRKPLLKAGIKLYELOPNHAVPAT 409
Db 348 LLATLKMTAQRGKIVDIIIP--KNDLSLMVQWASRAFYGELLEAGVQIHEF----- 396
Qy 410 KDKGLTSSVTSLSHAKTFIVDGKRIFIGSFNLDPRSARLNTMGVWIESPKIAEQM 465
Db 397 -DGGL-----LHTKSVVIDQQFCLVGTVDLDRSLWNLNFEFLTAVDDLEFTQOM 444
```


RESULT 32

AG0266
cardiolipin synthetase (EC 2.7.8.-) [imported] - Yersinia pestis (strain CO92)
C:Species: Yersinia pestis
C:Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 09-Jul-2004
C:Accession: AG0266
R:Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.;
deleno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;
11, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrall,
Nature 413, 523-527, 2001
A:Title: Genome sequence of Yersinia pestis, the causative agent of plague.
A:Reference number: AB0001; MUID:21470413; PMID:11586360
A:Accession: AG0266
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-486 <KUR>
A:Cross-references: UNIPROT:Q8ZEI2; UNIPARC:UPI00000CDBB3; GB:AL590842; PIDN:CAC90995.1;
C:Genetics:
A:Gene: cIs
C:Superfamily: cardiolipin synthase
C:Keywords: transferase

Query Match 8.9%; Score 241.5; DB 2; Length 486;
Best Local Similarity 22.5%; Pred. No. 1.5e-09;
Matches 112; Conservative 81; Mismatches 180; Indels 125; Gaps 23;

```
QY 33 SCSSWLPLEBTERSRH-ENTSKPVLDDNLIQIRHTPHN-NGL--SDIYLLDDPHEALAA 88
DB 77 STAKWLSELK---ECQIFANSNEVASPFCQRCRGGINGVKGNGQLQLTTTDDTLKA 133
QY 89 RAALIESAEHSLDQYYIWR-NDISGRLLFNLMYLAARGVVRVRLDDNNTR----- 140
DB 134 LVRLDIELARHNIEVFIWQPGGLVDQVAESLM-AAARRGVHCELLDLSAGSKQFFRSPY 192
QY 141 -----GLDLLLLALSDHPNIEVRLFPVLRKWRALGYLTDPPRLNRMRHNSFTADN 193
DB 193 PAMMRNAGIEVW-----EALKVNVFRMFL-----RRMDLRQHRKIVLIDN 232
QY 194 RATILGGRNIGDEYFKVGEDTVFADLDILA-----TGSVVGEVSHDFDRYWASHAHNAT 248
DB 233 YVAYTGSMNMVDPFRFFQDQAGVGQWIDMARMGEFVATTIGV-----YACDWEIETGK 286
QY 249 RII---RSGNIGKQLQALGYNDETSRHALLRYRTVEQSPLYQKIQTGRIDWQSVQRLI 305
DB 287 RILPPPDANI-----MPFEETG-HTI---QVIASGQFPE-----EMIHCALL 327
QY 306 SDTPAKGLDRDRKPPPIAGRLQDALKQPEKSVYLVSPYFVPTKSGTDALAKLVODGIDVT 365
DB 328 T-----AVYAAREQLIMTTPYFVPSDDLHLHAICTAAQRGVDVS 365
QY 366 VLTNLSQATDVAHVSGYVKYRKLKAGIKLYELQPNHAVPATKDKGLTGSSVTSIHLAK 425
DB 366 IIVP--RENDMMVRWASRAFFTELLNAGVKIYQF-----EGGL-----LHSK 406
QY 426 TFIVDGKRIFIGSNLDPRLNTEMGVVIESPKEIAEQMERTLADTPPEYAVRVTL--- 482
DB 407 SVLVGQLSLVGTVNLDMRSLNFEITLVIDD-----DGFADLAQVDDYIARSALLDG 462
QY 483 DKHN-LQWHDPAKTKTY 499
DB 463 ERWKRPLMHRVTERLFY 480
```

RESULT 33

S55419
cardiolipin synthetase homolog ywJ - Bacillus subtilis
C:Species: Bacillus subtilis
C:Date: 15-Jul-1995 #sequence_revision 01-Sep-1995 #text_change 09-Jul-2004
C:Accession: S55419; C70060
R:Glaser, P.; Danchin, A.
submitted to the EMBL Data Library, May 1995
A:Description: Cloning and sequencing of the Bacillus subtilis chromosomal region from 3

A:Reference number: S55414

A:Accession: S55419

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-398 <GUA>

A:Cross-references: UNIPROT:P45865; UNIPARC:UPI0000060BCA; EMBL:249782; NID:G853752; P1
R:Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Bert
C.; Bron, S.; Broutillet, S.; Brusch, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; C
A.; Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E
Nature 390, 249-256, 1997
A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gall
iech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.
Koetter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois
A:Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Mau
Y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.W.; Portetell
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadale, Y.; Sato, T.; Scanlon
A:Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, Y.; Sekowska, A.; Ser
akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpestra, P.; Tognoni, A.; Tosato, V.; Uchiyama
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida,
A:Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.
A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis
A:Reference number: A69580; MUID:98044033; PMID:9384377
A:Accession: C70060

A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA

A:Residues: 1-398 <KUN>

A:Cross-references: UNIPARC:UPI0000060BCA; GB:Z99123; GB:AL009126; NID:G2636240; PIDN:C

A:Experimental source: strain 168

C:Genetics:

A:Gene: ywJ

C:Superfamily: cardiolipin synthase

Query Match 8.6%; Score 235; DB 2; Length 398;

Best Local Similarity 20.7%; Pred. No. 3.3e-09;

Matches 93; Conservative 71; Mismatches 178; Indels 108; Gaps 12;

```
QY 74 SDIYLLDDPHEALAAARAALIESAEHSLDQYYIWRNDISGRLLFNLMYLAARGVVRVRL 133
DB 43 SDIELIHCGADIVRMWMDIRQAASSVHMFFIMKDEVSHNMVTLTKTKAQAGSVYLL 102
QY 134 LDNNTRTGLDLLLLALSDHPNIEVRLFN----PFVLRKWRALGYLTDPPRLNRMRHNSKF 189
DB 103 LDWAGCRAIKKTALQTMKNAGVHVVMNRPRFPFF-----FFHMQRNHRKIT 150
QY 190 TADNRATILGGRNIGDEYFKVGEDTVFADLDILATGVSVEVSHDFDRYWASHAHNATR 249
DB 151 VIDGKIYIGGFGFIABEY--LGKKAKFGNWEDYHL-RMIGEGVHDLQTLFASDLKRNT-- 205
QY 250 IIRSGNIGKQLQALGYNDETSRHALLRYRTVEQSPLYQKIQTGRI-----DWQSVQT 302
DB 206 -----GIE-LG-----SDVWPKLQOGTISHKIYATDGYSLN 236
QY 303 RLISDTPAKGLDRDRKPPPIAGRLQDALKQPEKSVYLVSPYFVPTKSGTDALAKLVODGI 362
DB 237 IYLAN-----IAQAKNRLTVCTPYIPSPKPLQEQALINARKNGV 274
QY 363 DVTVLNLSQATDVAHVSGYVKYRKLKAGIKLYELQPNHAVPATKDKGLTGSSVTSLS 422
DB 275 SVRIIVP--MKSDDPLVREAAFTYSELDDAGCLIIYYQGF-----Y 315
QY 423 HAKTFIVDGKRIFIGSNLDPRLNTEMGVVIESPKEIAEQMERTLADTPPEYAVRVTL 482
DB 316 HVKALLIIDDHLSIIGTANFDKSLFLNEEVNVEIDDEAFTSEVYALIEEDM----- 366
QY 483 DKHNRLQWHDPAKTKTYPNEPEAKLWKRIA 512
DB 367 -KKSELLTMEDFSRTFRQRPAAEWLGRALS 395
```

RESULT 34

AI0598

probable phospholipase ybho [imported] - Salmonella enterica subsp. enterica serovar Typh
C:Species: Salmonella enterica subsp. enterica serovar Typhi

[illegible]

Qy	264	GYNDETSRHALLRYRETVFQSPLYQKI--QTGRIDWQSVQTRILISDTPAKGLDRDRKPP	321
Db	283	-----SVRPDKKLEREDSYFVPPEEGNIDIQ-----IVANGPKS--DNKTLRTG	326
Qy	322	IAGRLDALQKPEKSVVLVSPYFVPVTKSGTDALAKLVQDGIDVTVLVTNSLQAATDVAAVHS	381
Db	327	FIKIWMDA---EDYIWLQSPYILIPDSSMITALVAAANGVDVRIIMPNM--PDHFPIPR	380
Qy	382	GYVKYRKPLKAGIKUYELQPNHAVPATKKGITGSGVTSLSHAKTPIVQGRKIFISFNL	441
Db	381	ATQYVANYLHKHGKVTINYTNFG-----IHSKTLVMDGKLGVFGTNQ	423
Qy	442	DPESARLNTGEMGVVIESPKIAEQMERT	468
Db	424	DIRSYELNFISAFCYDETVAKEMSKT	450
RESULT 36			
AC2882			
cardiolipin synthase [imported] - Agrobacterium tumefaciens (strain C58, Dupont			
C/Species: Agrobacterium tumefaciens			
C/Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 09-Jul-2004			
C/Accession: AC2882			
R/Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y			
erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutyaavin, T.; Levy, R.; Li, M.			
; Karp, P.; Romero, P.; Zhang, S.			
Science 294, 2317-2323, 2001			
A/Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gord			
ster, E.W.			
A/Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens			
A/Reference number: AB5577; MUID:21608550; PMID:11743193			

```

A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-492 <KUR>
A:Cross-references: UNIPROT:Q8UCJ8; UNIPARC:UPI000000D1F08; GB:AE008688; PIDN:AA
A:Experimental source: strain C58 (Dupont)
C:Genetics:
A:Gene: cjs
A:Map position: circular chromosome
C:Superfamily: cardiolipin synthase

Query Match      8.3%  Score 224.5; DB 2; Length 492;
Best Local Similarity 23.0%; Pred. No. 2.6e-08;
Matches 96; Conservative 56; Mismatches 190; Indels 75; Gaps 11;

Qy      48 RHFNTSKPVLDNLIQIRHT--PHNNGLSDTYLLDDPHEALAAARAALIESAEHSLDLQYYI 106
Db      106 RFEAMK-TLGRDV--TRHALTTGNGIEPLVSGD--VAYAAMLEAIGEAKRISIILETYI 159

Qy      107 WRNDISGRLLFNLMYLAERGCVRVRLDNDNTTGLDLLLLALDSDHPNIEVRLFNFPVLR 166
Db      160 FDNDRTGAREVAALERAKRGVVRVLIDAVGARYSVPSILPTLRDKDIVADVFNGNVIM 219

```

```
Qy 167 KWRALGYLTDPRNRRMHKSFADNRATILGGRNIGDEY-FKVGEDTVFADLDILATG 225
Db 220 GLR-----LPYANLRTHRKILVVDGRIAPSGMNIREGFTLEFGESQSHDTHFKITG 272
Qy 226 SVGGEVSHDFRYWASHAHNATRIIRSGNIGKGLQALGYNDTSRHALLRYRTVEQSP 285
Db 273 PVSDFFSIAAEDWRFTTGE-----VLDAP 297
Qy 286 LYQKIOTGRIDWQSVQTRLSIDTPAKGLDRDRKRPPIAGRLQDALKOPEKSVLVSPYFV 345
Db 298 VMDIAPDGVPGSQIARVCSSGPDKSIETSHK-----MLMGAFVSARSILLIMPYFL 351
Qy 346 PTKSGTDALAKLVQDGDIVTVLTNSLOATDVAAVHSGYVKYRKPLKAGIKLYELOPNHA 405
Db 352 PDRELISALITAARRGCVVDIIVP--KSNNLVLDRAWTAQPDQMLKNYCIWR----- 403
Qy 406 VPATKDKLGTGSSVTSLSHAKTFIVDGKRIFGSNLDPKRSARLNTMGVVIESPZIA 462
Db 404 -----ATGA-----FNHSKLLVIDGRWSYIGSSNLDPRSLRNLFEIDLEVMDEEFA 449

RESULT 37
B97658
cardiolipin synthase (PA5394) [imported] - Agrobacterium tumefaciens (strain C58, Cereon
C;Species: Agrobacterium tumefaciens
C;Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 09-Jul-2004
C;Accession: B97658
A;: Liu, F.; Wollam, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldman,
R.; Goodner, B.; Hinkle, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.;
Science 294, 2323-2328, 2001
A;Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum
A;Reference number: A97359; MUID:21608551; PMID:11743194
A;Accession: B97658
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-492 <KUR>
A;Cross-references: UNIPROT:Q8UCJ8; UNIPARC:UPI00000DIF08; GB:AE007869; PIDN:AAK88219.1;
C;Genetics:
A;Gene: AGR C 4515
A;Map position: circular chromosome
C;Superfamily: cardiolipin synthase

Query Match 8.3%; Score 224.5; DB 2; Length 492;
Best Local Similarity 23.0%; Pred. No. 2.6e-08;
Matches 96; Conservative 56; Mismatches 190; Indels 75; Gaps 11;

Qy 48 RHFNTPKPVLLDNLQIRHT-PHNNGLSDIYLLDDPHEALAAARALIESAHSLSLDQYII 106
Db 106 RRFEAMK-TLGDV--TRHALTTGNGIEPLVSGD---VAYANMLEAIGEAKRSIILETYI 159
Qy 107 WRNDISGRLLFNLMYLAARGVVRVRLDDNNTRGLDLDLALDSDHNPNIENVLPFVLR 166
Db 160 FDNDRGARFVAALERAKLGEVVRVLDIVAGARYSVPSLPTLRDKDIVADVFNQNVIM 219
Qy 167 KWRALGYLTDPRNRRMHKSFADNRATILGGRNIGDEY-FKVGEDTVFADLDILATG 225
Db 220 GLR-----LPYANLRTHRKILVVDGRIAPSGMNIREGFTLEFGESQSHDTHFKITG 272
Qy 226 SVGGEVSHDFRYWASHAHNATRIIRSGNIGKGLQALGYNDTSRHALLRYRTVEQSP 285
Db 273 PVSDFFSIAAEDWRFTTGE-----VLDAP 297
Qy 286 LYQKIOTGRIDWQSVQTRLSIDTPAKGLDRDRKRPPIAGRLQDALKOPEKSVLVSPYFV 345
Db 298 VMDIAPDGVPGSQIARVCSSGPDKSIETSHK-----MLMGAFVSARSILLIMPYFL 351
Qy 346 PTKSGTDALAKLVQDGDIVTVLTNSLOATDVAAVHSGYVKYRKPLKAGIKLYELOPNHA 405
Db 352 PDRELISALITAARRGCVVDIIVP--KSNNLVLDRAWTAQPDQMLKNYCIWR----- 403
Qy 406 VPATKDKLGTGSSVTSLSHAKTFIVDGKRIFGSNLDPKRSARLNTMGVVIESPZIA 462
Db 404 -----ATGA-----FNHSKLLVIDGRWSYIGSSNLDPRSLRNLFEIDLEVMDEEFA 449
```

RESULT 38

B83377

probable phospholipase PA2155 [imported] - Pseudomonas aeruginosa (strain PA01)

C;Species: Pseudomonas aeruginosa

C;Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2004

C;Accession: B83377

R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; B

adman, S.; Yuan, Y.; Brody, L.B.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Li

; Lory, S.; Olson, M.V.

Nature 406, 959-964, 2000

A;Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic path

A;Reference number: A82950; MUID:20437337; PMID:10584043

A;Accession: B83377

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-401 <STO>

A;Cross-references: UNIPROT:Q911W0; UNIPARC:UPI00000C55D9; GB:AE004642; GB:AE004091; NI

A;Experimental source: strain PA01

C;Genetics:

A;Gene: PA2155

C;Superfamily: cardiolipin synthase

Query Match 8.2%; Score 224; DB 2; Length 401;

Best Local Similarity 22.8%; Pred. No. 2.1e-08;

Matches 98; Conservative 59; Mismatches 184; Indels 88; Gaps 11;

Qy 68 PHNGLSDIYLLDDPHEALAAARALIESAHSLSLDQYIWRNDISGRLLFNLMYLAARG 127

Db 4 PWRDG-NRVELLNGEYFPRLPQCIARREILLETFFIFEDEVGRQLQEALSAARERG 62

Qy 128 VYVRLDDNNTRGLDLDLALDSDHNPNIENVLPFVLRKWRALGYLTDPRNRRMHK 187

Db 63 VEQVTVDGYGTASLSPDYLRATASGVRVHLFDP----KPRLLGWRTN---LFRRLHRK 115

Qy 188 SFTADNRATILGGRNIGDEYFKVGEDTVFADLDILATGSVVGESHDFDRYWASHAHNA 247

Db 116 LVVIDRRQAFVGGINVGEDH----- 135

Qy 248 TRIIRSGNIGK---GLQALGYNDTSRHALLRYRTVEQSPLYQKIQTGRIDWQSVQTRL 304

Db 136 --LVRRGNMAKQDYAVRVEGPPVVRDIRQACLALLEDADYPPPLRPSGAG---QPARVRL 189

Qy 305 ISDTPAKGLDRDRKRPPIAGRLQDALKOPEKSVLVSPYFVPTKSGTDALAKLVQDGDIV 364

Db 190 VI-----RNDQSSDDIERYLQAIRQARRRLLIANAYFFPGYRLRLRLRDAARGVRV 243

Qy 365 TVLTNSLOATDVAAVHSGYV-KYRKPLLKAGIKLYEL--QPNHAVPATKDKGLTGSSVTS 421

Db 244 DLVLQGMPPMPLVRLCSRLLYDY---LLREGVRIHYCQRP----- 281

Qy 422 LHAKTIVDGKRIFGSNLDPKRSARLNTMGVVIESPZIAEQMERTLADTTPEYAYRVT 481

Db 282 LHGKVAVIDDWSITGSSNLDPLSLNLEANLVIRDVAFNGQLYQHLRELARRHCRIS 341

Qy 482 LDKHNEFLQW 490

Db 342 RRHARRGYW 350

RESULT 39

E64815

Ybho protein - Escherichia coli (strain K-12)

N;Alternate names: protein b0789

C;Species: Escherichia coli

C;Date: 12-Sep-1997 #sequence_revision 17-Sep-1997 #text_change 09-Jul-2004

C;Accession: E64815

R;Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; C

A.; Rose, D.J.; Mau, B.; Shao, Y.

Science 277, 1453-1462, 1997

A;Title: The complete genome sequence of Escherichia coli K-12.

A;Reference number: A64720; MUID:97426617; PMID:9278503

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: May 2, 2006, 06:06:00 ; Search time 187 Seconds
(without alignments)
1233.550 Million cell updates/sec

Title: US-10-665-990A-14

Perfect score: 2720

Sequence: 1 MHTDPKIQAMSETISPMKT.....KLWKRIAAKILSLPIEGLL 525

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_21:*

- 1: Geneseqp1980s:*
- 2: Geneseqp1990s:*
- 3: Geneseqp2000s:*
- 4: Geneseqp2001s:*
- 5: Geneseqp2002s:*
- 6: Geneseqp2003as:*
- 7: Geneseqp2003bs:*
- 8: Geneseqp2004s:*
- 9: Geneseqp2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2712	99.7	525	6	ABP79466 N. gonorr
2	2643	97.2	525	5	Abb78070 Amino aci
3	2606.5	95.8	507	3	AAY75751 Neisseria
4	2586	95.1	508	3	AAY75753 Neisseria
5	2573	94.6	508	3	AAY75752 Neisseria
6	831.5	30.6	543	8	ADL04719 M. catarr
7	790	29.0	564	7	ABO75163 Pseudomon
8	632	23.2	502	2	AAW55452 H. pylori
9	632	23.2	502	2	AAW55550 H. pylori
10	590	21.7	428	2	AAW98378 H. pylori
11	493.5	18.1	271	6	ADA34272 Acinetoba
12	478.5	17.6	275	6	ADA34304 Acinetoba
13	419	15.4	2519	7	ABM88218 Rice abio
14	409	15.0	321	6	ADA36398 Acinetoba
15	308	11.3	513	6	ABM71816 Staphyloc
16	296	10.9	486	7	ADG30702 Xanthomon
17	294.5	10.8	504	5	ABBA49813 Listeria
18	291	10.7	495	8	ABP39428 Staphyloc
19	291	10.7	495	8	ABP39428 Staphyloc
20	287	10.6	482	5	ABBA49993 Listeria
21	286.5	10.5	493	5	ABB32481 Staphyloc
22	286.5	10.5	493	7	ADC21336 Staphyloc
23	286.5	10.5	493	7	ADC25022 S. aureus
24	286.5	10.5	493	7	ADD52524 Staphyloc

25	286.5	10.5	502	5	ABB32493 Staphyloc
26	286.5	10.5	502	7	ADC21402 Staphyloc
27	286.5	10.5	502	7	ADC25088 S. aureus
28	286.5	10.5	502	7	ADD52590 Staphyloc
29	286.5	10.5	510	6	ABM71594 Staphyloc
30	282.5	10.4	480	5	ABB54276 Lactococc
31	275	10.1	420	9	ABM97264 M. xanthu
32	274.5	10.1	206	2	AAW55249 H. pylori
33	274.5	10.1	490	6	ABM69020 Photorhab
34	268.5	9.9	584	7	ABO75565 Pseudomon
35	268	9.9	579	7	ABO68379 Pseudomon
36	259.5	9.5	494	5	ABP39331 Staphyloc
37	259.5	9.5	494	8	ADSO5084 Staphyloc
38	252	9.3	166	6	ADA36341 Acinetoba
39	252	9.3	422	4	ABM76548 Corynebac
40	252	9.3	500	4	AG92760 C glutami
41	252	9.3	500	5	AG80237 C. glutam
42	246	9.0	508	7	ADF04432 Bacterial
43	241	8.9	409	7	ADH85779 Enterococ
44	232	8.5	442	4	AG81953 S. epider
45	231.5	8.5	386	7	ADC94461 E. faeciu

ALIGNMENTS

RESULT 1

ABP79466

ID ABP79466 standard; protein; 525 AA.

XX AC ABP79466;

XX AC

DT 07-MAR-2003 (first entry)

XX DE N. gonorrhoeae amino acid sequence SEQ ID 5462.

XX KW Antibacterial; infection; vaccine; gene therapy.

XX OS Neisseria gonorrhoeae.

XX PN WO200279243-A2.

XX PD 10-OCT-2002.

PF 12-FEB-2002; 2002WO-IB002069.

PR 12-FEB-2001; 2001GB-00003424.

XX (CHIR-) CHIRON SPA.

XX Fontana MR, Pizza M, Masignani V, Monaci E;

DR WPI; 2003-058415/05.

XX N-PSDB; ABZ40436.

PT New protein from Neisseria gonorrhoeae, useful for the manufacture of a medicament for treating or preventing N. gonorrhoeae infection.

XX Disclosure; Page 587; 815pp; English.

XX The present invention relates to proteins from Neisseria gonorrhoeae. Also disclosed are the nucleic acid molecules encoding the proteins and antibodies that specifically bind to the proteins. The composition comprising the protein, nucleic acid or antibody is useful for the manufacture of a medicament for treating or preventing N. gonorrhoeae infection, this may be in the form of a vaccine or gene therapy.

XX Sequences given in records ABP76736-ABP81046 represent nucleic acid molecules of the invention

XX Sequence 525 AA;

Query Match 99.7%; Score 2712; DB 6; Length 525;

Best Local Similarity 99.6%; Pred. No. 9.9e-240;

Matches	523	Conservative	1	Mismatches	1	Indels	0	Gaps	0
Qy	1	MHTDPKIQAMPSETISPMKTRSLISLCLLCSCSSMLPPEERTESRHFTSKPVL	LDN	60					
Db	1	MHTDPKIQAMPSETISPMKTRSLISLCLLCSCSSMLPPEERTESRHFTSKPVL	LDN	60					
Qy	61	ILQIRHTPHNNGLSDIYLLDDPHBALAARAALIESAHSLDQYIWRNDISGRLL	FNLM	120					
Db	61	ILQIRHTPHNNGLSDIYLLDDPHBAFAARAALIESAHSLDQYIWRNDISGRLL	FNLM	120					
Qy	121	YLAAGRVRRVRLLLDDNNTGRLDLLLALDSHPNIEVRLNFPVLRKWRALGYLT	DPRL	180					
Db	121	YLAAGRVRRVRLLLDDNNTGRLDLLLALDSHPNIEVRLNFPVLRKWRALGYLT	DPRL	180					
Qy	181	NRRMHKNSFTADNRATILGGRNIGDEYFKVGEDTVFADLDILATGSSVVGESH	DFR	240					
Db	181	NRRMHKNSFTADNRATILGGRNIGDEYFKVGEDTVFADLDILATGSSVVGESH	DFR	240					
Qy	241	SHSAHNATRIIRSGNIGKGLQALGYNDETSRHALLRYRETVQESPLYOKIOT	GRIDWQSV	300					
Db	241	SHSAHNATRIIRSGNIGKGLQALGYNDETSRHALLRYRETVQESPLYOKIOT	GRIDWQSV	300					
Qy	301	QTRLISTPAKGLDRDRKPPIAGRLQDALQKPEKSVYLVSPYFVPTKSGTDALAK	LQVD	360					
Db	301	QTRLISDSPAKGLDRDRKPPIAGRLQDALQKPEKSVYLVSPYFVPTKSGTDALAK	LQVD	360					
Qy	361	GIDVTVLNTSLQATDVAHVSGYVKYRPLLKAGIKLYELQPNHAVPAKDKGLTG	SSVT	420					
Db	361	GIDVTVLNTSLQATDVAHVSGYVKYRPLLKAGIKLYELQPNHAVPAKDKGLTG	SSVT	420					
Qy	421	SLHAKTFIVDGKRIFIGSFNLDPRSARLNTMGVVIESPKLAEQMERTLADTTPE	YARV	480					
Db	421	SLHAKTFIVDGKRIFIGSFNLDPRSARLNTMGVVIESPKLAEQMERTLADTTPE	YARV	480					
Qy	481	TLDKHNLQWHDPAATRTKYPNPEPAKUWKRIIAAKILSLLP	IEGLL	525					
Db	481	TLDKHNLQWHDPAATRTKYPNPEPAKUWKRIIAAKILSLLP	IEGLL	525					

RESULT 2
ABB78070
ID ABB78070 standard; protein: 525 AA.

[illegible]

DE Amino acid sequence of p55 polypeptide.

AA p177; p88; p64; p55; p46; p33; p30; p29; p28; p27; p26; p25; p24; p23; p22; p21; p20; p19; p18; p17; p16; p15; p14; p13; p12; p11; p10; p9; p8; p7; p6; p5; p4; p3; p2; p1; p0; p-1; p-2; p-3; p-4; p-5; p-6; p-7; p-8; p-9; p-10; p-11; p-12; p-13; p-14; p-15; p-16; p-17; p-18; p-19; p-20; p-21; p-22; p-23; p-24; p-25; p-26; p-27; p-28; p-29; p-30; p-31; p-32; p-33; p-34; p-35; p-36; p-37; p-38; p-39; p-40; p-41; p-42; p-43; p-44; p-45; p-46; p-47; p-48; p-49; p-50; p-51; p-52; p-53; p-54; p-55; p-56; p-57; p-58; p-59; p-60; p-61; p-62; p-63; p-64; p-65; p-66; p-67; p-68; p-69; p-70; p-71; p-72; p-73; p-74; p-75; p-76; p-77; p-78; p-79; p-80; p-81; p-82; p-83; p-84; p-85; p-86; p-87; p-88; p-89; p-90; p-91; p-92; p-93; p-94; p-95; p-96; p-97; p-98; p-99; p-100; p-101; p-102; p-103; p-104; p-105; p-106; p-107; p-108; p-109; p-110; p-111; p-112; p-113; p-114; p-115; p-116; p-117; p-118; p-119; p-120; p-121; p-122; p-123; p-124; p-125; p-126; p-127; p-128; p-129; p-130; p-131; p-132; p-133; p-134; p-135; p-136; p-137; p-138; p-139; p-140; p-141; p-142; p-143; p-144; p-145; p-146; p-147; p-148; p-149; p-150; p-151; p-152; p-153; p-154; p-155; p-156; p-157; p-158; p-159; p-160; p-161; p-162; p-163; p-164; p-165; p-166; p-167; p-168; p-169; p-170; p-171; p-172; p-173; p-174; p-175; p-176; p-177; p-178; p-179; p-180; p-181; p-182; p-183; p-184; p-185; p-186; p-187; p-188; p-189; p-190; p-191; p-192; p-193; p-194; p-195; p-196; p-197; p-198; p-199; p-200; p-201; p-202; p-203; p-204; p-205; p-206; p-207; p-208; p-209; p-210; p-211; p-212; p-213; p-214; p-215; p-216; p-217; p-218; p-219; p-220; p-221; p-222; p-223; p-224; p-225; p-226; p-227; p-228; p-229; p-230; p-231; p-232; p-233; p-234; p-235; p-236; p-237; p-238; p-239; p-240; p-241; p-242; p-243; p-244; p-245; p-246; p-247; p-248; p-249; p-250; p-251; p-252; p-253; p-254; p-255; p-256; p-257; p-258; p-259; p-260; p-261; p-262; p-263; p-264; p-265; p-266; p-267; p-268; p-269; p-270; p-271; p-272; p-273; p-274; p-275; p-276; p-277; p-278; p-279; p-280; p-281; p-282; p-283; p-284; p-285; p-286; p-287; p-288; p-289; p-290; p-291; p-292; p-293; p-294; p-295; p-296; p-297; p-298; p-299; p-300; p-301; p-302; p-303; p-304; p-305; p-306; p-307; p-308; p-309; p-310; p-311; p-312; p-313; p-314; p-315; p-316; p-317; p-318; p-319; p-320; p-321; p-322; p-323; p-324; p-325; p-326; p-327; p-328; p-329; p-330; p-331; p-332; p-333; p-334; p-335; p-336; p-337; p-338; p-339; p-340; p-341; p-342; p-343; p-344; p-345; p-346; p-347; p-348; p-349; p-350; p-351; p-352; p-353; p-354; p-355; p-356; p-357; p-358; p-359; p-360; p-361; p-362; p-363; p-364; p-365; p-366; p-367; p-368; p-369; p-370; p-371; p-372; p-373; p-374; p-375; p-376; p-377; p-378; p-379; p-380; p-381; p-382; p-383; p-384; p-385; p-386; p-387; p-388; p-389; p-390; p-391; p-392; p-393; p-394; p-395; p-396; p-397; p-398; p-399; p-400; p-401; p-402; p-403; p-404; p-405; p-406; p-407; p-408; p-409; p-410; p-411; p-412; p-413; p-414; p-415; p-416; p-417; p-418; p-419; p-420; p-421; p-422; p-423; p-424; p-425; p-426; p-427; p-428; p-429; p-430; p-431; p-432; p-433; p-434; p-435; p-436; p-437; p-438; p-439; p-440; p-441; p-442; p-443; p-444; p-445; p-446; p-447; p-448; p-449; p-450; p-451; p-452; p-453; p-454; p-455; p-456; p-457; p-458; p-459; p-460; p-461; p-462; p-463; p-464; p-465; p-466; p-467; p-468; p-469; p-470; p-471; p-472; p-473; p-474; p-475; p-476; p-477; p-478; p-479; p-480; p-481; p-482; p-483; p-484; p-485; p-486; p-487; p-488; p-489; p-490; p-491; p-492; p-493; p-494; p-495; p-496; p-497; p-498; p-499; p-500; p-501; p-502; p-503; p-504; p-505; p-506; p-507; p-508; p-509; p-510; p-511; p-512; p-513; p-514; p-515; p-516; p-517; p-518; p-519; p-520; p-521; p-522; p-523; p-524; p-525; p-526; p-527; p-528; p-529; p-530; p-531; p-532; p-533; p-534; p-535; p-536; p-537; p-538; p-539; p-540; p-541; p-542; p-543; p-544; p-545; p-546; p-547; p-548; p-549; p-550; p-551; p-552; p-553; p-554; p-555; p-556; p-557; p-558; p-559; p-560; p-561; p-562; p-563; p-564; p-565; p-566; p-567; p-568; p-569; p-570; p-571; p-572; p-573; p-574; p-575; p-576; p-577; p-578; p-579; p-580; p-581; p-582; p-583; p-584; p-585; p-586; p-587; p-588; p-589; p-590; p-591; p-592; p-593; p-594; p-595; p-596; p-597; p-598; p-599; p-600; p-601; p-602; p-603; p-604; p-605; p-606; p-607; p-608; p-609; p-610; p-611; p-612; p-613; p-614; p-615; p-616; p-617; p-618; p-619; p-620; p-621; p-622; p-623; p-624; p-625; p-626; p-627; p-628; p-629; p-630; p-631; p-632; p-633; p-634; p-635; p-636; p-637; p-638; p-639; p-640; p-641; p-642; p-643; p-644; p-645; p-646; p-647; p-648; p-649; p-650; p-651; p-652; p-653; p-654; p-655; p-656; p-657; p-658; p-659; p-660; p-661; p-662; p-663; p-664; p-665; p-666; p-667; p-668; p-669; p-670; p-671; p-672; p-673; p-674; p-675; p-676; p-677;

OS *Neisseria gonorrhoeae*.

PN WO200260936-A2.

08-AUG-2002.

31-JAN-2002: 2002WO-US002881.

31-JAN-2001; 2001US-0266070P.

PR 23-OCT-2001; 2001US-0344452P.
PR 23-OCT-2001; 2001US-0344452P.

PA (IOWA) UNIV IOWA RES FOUNO.

PA (APIC/) APICELLA M A.

PA (GIBS/) GIBSON B W.

PA (BROW/) BROWN E.

PI Apicella MA, Edwards

Apicella MA, Edwards JL, Gibson BW, Scheffler K, Brown E;

XX	WFI; 2002-619227/66.
DR	N-PSDB; ABQ78301.
DR	
XX	
PT	New polypeptide comprising p177, p88, p64, p35 or p46 from Neisseria
PT	gonorrhoeae, useful for preventing, or protecting a female patient
PT	against, N. gonorrhoeae colonization or infection.
XX	
XX	Claim 7; Page 120-121; 130pp; English.
XX	
CC	The present sequence represents a p55 polypeptide. The specification
CC	describes p177, p88, p64, p35 and p46 polypeptides from Neisseria
CC	gonorrhoeae. The polypeptides are useful as vaccines, for preventing, or
CC	protecting a female patient against, N. gonorrhoeae colonization or
CC	infection. Such immunisation can prevent gonorrhoea in women. (Updated on
CC	29-AUG-2003 to standardise OS field)
XX	
XX	Sequence 525 AA;
SQ	

Query Match	97.2%;	Score 2643;	DB 5;	Length 525;
Best Local Similarity	97.3%;	Pred. No. 2.2e-233;		
Matches 511: Conservative		5; Mismatches	9;	Indels

RESULT 3
AAY75751
ID AAY75751 standard; protein: 507 AA.

XX
DT 12-SEP-2003 (revised)21-MAR-2000 (first entry)
XX

DE
XX
NETSSETA gonnimeae OKF 50, process sequence 55Q ID NO. 2574

Neisseria meningitidis; neisseria gonorrhoeae; antigen; vaccine;
KW antigenic: diagnosis; immunogenic; infection; meningitis; septicaemia;
KW

KW antibacterial; gene therapy.
 XX
 OS Neisseria gonorrhoeae.
 XX
 PN WO9957280-A2.
 XX
 PD 11-NOV-1999.
 XX
 PF 30-APR-1999; 99WO-US009346.
 XX
 PR 01-MAY-1998; 98US-0083758P.
 PR 31-JUL-1998; 98US-0094869P.
 PR 02-SEP-1998; 98US-0098994P.
 PR 02-SEP-1998; 98US-0099062P.
 PR 09-OCT-1998; 98US-0103749P.
 PR 09-OCT-1998; 98US-0103749P.
 PR 09-OCT-1998; 98US-0103796P.
 PR 25-FEB-1999; 99US-0121528P.
 XX
 PA (CHIR) CHIRON CORP.
 PA (GENO-) INST GENOMIC RES.
 XX
 PI Fraser C, Galeotti C, Grandi G, Hickey E, Massignani V, Mora M;
 PI Petersen J, Pizsa M, Rappuoli R, Ratti G, Scalato E, Scarselli M;
 PI Tettelin H, Venter JC;
 XX
 DR WPI; 2000-062150/05.
 DR N-PSDB; AAZ54513.
 XX
 PT Novel Neisserial polypeptides predicted to be useful antigens for
 PT vaccines and diagnostics.
 XX
 PS Claim 2; Page 1394; 1453pp; English.
 XX
 CC AAZ53015 to AAZ54536, AAZ54577 to AAZ54615, and AAZ74253 to AAZ75941
 CC represent novel Neisseria meningitidis and N. gonorrhoeae polynucleotides
 CC and polypeptides. AAZ54537 to AAZ54576 and AAZ54616 to AAZ5473 represent
 CC PCR primers used in the exemplification of the present invention. The
 CC polypeptides, the polynucleotides, antibodies and compositions of the
 CC invention can be used as vaccines, as diagnostic reagents, and as
 CC immunogenic compositions. The polypeptides can be used in the manufacture
 CC of medicaments for treating or preventing infection due to Neisserial
 CC bacteria (e.g. meningitis and septicemia), to detect the presence of
 CC Neisseria bacteria, or to raise antibodies. They may also be used to
 CC screen for agonists or antagonists, which may themselves have use as
 CC antibacterial agents. The polynucleotides of the invention may also be
 CC used in gene therapy protocols. (Updated on 12-SEP-2003 to standardise OS
 CC field)
 XX
 SQ Sequence 507 AA;

Query Match 95.8%; Score 2606.5; DB 3; Length 507;
 Best Local Similarity 99.4%; Pred. No. 4.6e-230;
 Matches 508; Conservative 1; Mismatches 1; Indels 1; Gaps 1;
 18 MKTRSLISLLCLLSCSSWLPPLLEERTESRHFNTSKPVLDDNLIQIRHTPHNNGLSDIY 77
 1 MKTRSLISLLCLLSCSSWLPPLLEERTESRHFNTSKPVLDDNLIQIRHTPHNNGLSDIY 60
 78 LDDPHEALAAALIESAHSLSLDQYIWRNDISGRLLFNLMYLAAGRGVRVRLLDN 137
 61 LDDPHEAFARAALIESAHSLSLDQYIWRNDISGRLLFNLMYLAAGRGVRVRLLDN 120
 138 NTRGLDLLLLDSDHPNIEVRLNPFVLRKWRALGYLTDFPRLNRRMHNKSFADNPRATI 197
 121 NTRGLDLLLLDSDHPNI - VRLNPFVLRKWRALGYLTDFPRLNRRMHNKSFADNPRATI 179
 198 LGGRNIGDEYFKVGEDTVFADLDLILATGCVVGEVSHDFDRYWASHSAHNATRIIRSGNIG 257
 180 LGGRNIGDEYFKVGEDTVFADLDLILATGCVVGEVSHDFDRYWASHSAHNATRIIRSGNIG 239
 258 KGLQALGYNDETSRHALLRYRETVEQSPLYQKIQTGRIDWQSVQTRLISDSPAKGLDRDR 317

Db 240 KGLQALGYNDETSRHALLRYRETVEQSPLYQKIQTGRIDWQSVQTRLISDSPAKGLDRDR 299
 Qy 318 RKPIAGRLQDALKQPEKSVYLYSPYFVPTKSGTDALAKLVQDGDVTVLTNSLQATDVA 377
 Db 300 RKPIAGRLQDALKQPEKSVYLYSPYFVPTKSGTDALAKLVQDGDVTVLTNSLQATDVA 359
 Qy 378 AVHSGYVYKRPKLLKAGIKLYELQPNHAVPATKDKGLTGSSTVLSLHAKTFIVDGKRIFIG 437
 Db 360 AVHSGYVYKRPKLLKAGIKLYELQPNHAVPATKDKGLTGSSTVLSLHAKTFIVDGKRIFIG 419
 Qy 438 SFNLDPRLSARLNTMGVVIESPKEIAEQMERTLADTTPEYAYRVYVTLDKHNLQWHDPAIRK 497
 Db 420 SFNLDPRLSARLNTMGVVIESPKEIAEQMERTLADTTPEYAYRVYVTLDKHNLQWHDPAIRK 479
 Qy 498 TYPNEPEAKLWKRIAAKILSLPIEGILL 525
 Db 480 TYPNEPEAKLWKRIAAKILSLPIEGILL 507

RESULT 4
 AAY75753
 ID AAY75753 standard; protein; 508 AA.
 XX
 AC AAY75753;
 XX
 DT 21-MAR-2000 (first entry)
 XX
 DE Neisseria meningitidis ORF 987 protein sequence SEQ ID NO:2978.
 XX
 KW Neisseria meningitidis; Neisseria gonorrhoeae; antigen; vaccine;
 KW antigenic; diagnosis; immunogenic; infection; meningitis; septicemia;
 KW antibacterial; gene therapy.
 XX
 OS Neisseria meningitidis.
 XX
 PN WO9957280-A2.
 XX
 PD 11-NOV-1999.
 XX
 PF 30-APR-1999; 99WO-US009346.
 XX
 PR 01-MAY-1998; 98US-0083758P.
 PR 31-JUL-1998; 98US-0094869P.
 PR 02-SEP-1998; 98US-0098994P.
 PR 02-SEP-1998; 98US-0099062P.
 PR 09-OCT-1998; 98US-0103749P.
 PR 09-OCT-1998; 98US-0103796P.
 PR 09-OCT-1998; 98US-0103796P.
 PR 25-FEB-1999; 99US-0121528P.
 XX
 PA (CHIR) CHIRON CORP.
 PA (GENO-) INST GENOMIC RES.
 XX
 PI Fraser C, Galeotti C, Grandi G, Hickey E, Massignani V, Mora M;
 PI Petersen J, Pizsa M, Rappuoli R, Ratti G, Scalato E, Scarselli M;
 PI Tettelin H, Venter JC;
 XX
 DR WPI; 2000-062150/05.
 DR N-PSDB; AAZ54515.
 XX
 PT Novel Neisserial polypeptides predicted to be useful antigens for
 PT vaccines and diagnostics.
 XX
 PS Claim 2; Page 1396-1397; 1453pp; English.
 XX
 CC AAZ53015 to AAZ54536, AAZ54577 to AAZ54615, and AAZ74253 to AAZ75941
 CC represent novel Neisseria meningitidis and N. gonorrhoeae polynucleotides
 CC and polypeptides. AAZ54537 to AAZ54576 and AAZ54616 to AAZ5473 represent
 CC PCR primers used in the exemplification of the present invention. The
 CC polypeptides, the polynucleotides, antibodies and compositions of the
 CC invention can be used as vaccines, as diagnostic reagents, and as
 CC immunogenic compositions. The polypeptides can be used in the manufacture
 CC of medicaments for treating or preventing infection due to Neisserial
 CC

```
CC bacteria (e.g. meningitis and septicemia), to detect the presence of
CC Neisseria bacteria, or to raise antibodies. They may also be used to
CC screen for agonists or antagonists, which may themselves have use as
CC antibacterial agents. The polynucleotides of the invention may also be
CC used in gene therapy protocols
XX
SQ Sequence 508 AA;

Query Match          95.1%; Score 2586; DB 3; Length 508;
Best Local Similarity 98.2%; Pred. No. 3.5e-228;
Matches 499; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 18 MKTRSLSLLCLLCCSSWLPPLLEERTESRHFNTPKPVLLDNLILQIRHTPHNGLSDIY 77
DB 1 MKTRSLSLLCLLCCSSWLPPLLEERTESRHFNTPKPVLLDNLILQIRHTPHNGLSDIY 60

QY 78 LLDDPHEAALARAALIESAHSLSLQYYIWRNDISGRLLFNMLYLAARGVVRVLLDDN 137
DB 61 LLNDPHEAFAARAALIESAHSLSLQYYIWRNDISGRLLFNMLYLAARGVVRVLLDDN 120

QY 138 NTRGLDDLLLALDSDHPNIEVRLFNPFVLRKWRALGYLTFPRLNRRMNKSFADNRATI 197
DB 121 NTRGLDDLLLALDSDHPNIEVRLFNPFVLRKWRALGYLTFPRLNRRMNKSFADNRATI 180

QY 198 LGGRNIGDEYFKVGEDTVFADLDIATGVSVEVSHDFDRYWASHAHNATRIIRSGNIG 257
DB 181 LGGRNIGDEYFKVGEDTVFADLDIATGVSVEVSHDFDRYWASHAHNATRIIRSGNIG 240

QY 258 KGLQALGYNDTSRHALLRYRETVEQSPLYQKIQTGRIDWQSVQRLISDTPAKGLDRDR 317
DB 241 KGLQALGYNDTSRHALLRYRETVEQSPLYQKIQTGRIDWQSVQRLISDTPAKGLDRDR 300

QY 318 RKPPIAGRLQDALKQPEKSVLVSPYFVPTKSGTDALAKLVQDGDIVTVLNSLOATDVA 377
DB 301 RKPPIAGRLQDALKQPEKSVLVSPYFVPTKSGTDALAKLVQDGDIVTVLNSLOATDVA 360

QY 378 AVHSGYVKYRKPPLKAGIKLYELOPNHAVPATKDKGLTGSSVTSLHAKTFIVDGKRIFIG 437
DB 361 AVHSGYVKYRKPPLKAGIKLYELOPNHAVPATKDKGLTGSSVTSLHAKTFIVDGKRIFIG 420

QY 438 SFNLDPRLSARLNTMGVVIESPKEIAEQMERTLADTTPPEYAVRVTLTKENRLQWHPDPAK 497
DB 421 SFNLDPRLSARLNTMGVVIESPKEIAEQMERTLADTTPPEYAVRVTLTKENRLQWHPDPAK 480

RESULT 5
AAY75752
ID AAY75752 standard; protein; 508 AA.
AC
XX
XX AAY75752;
XX
DT 21-MAR-2000 (first entry)
XX
DE Neisseria meningitidis ORF 987 protein sequence SEQ ID NO:2976.
XX
KW Neisseria meningitidis; Neisseria gonorrhoeae; antigen; vaccine;
KW antigenic; diagnosis; immunogenic; infection; meningitis; septicemia;
KW antibacterial; gene therapy.
XX
OS Neisseria meningitidis.
XX
XX WO9957280-A2.
XX
XX 11-NOV-1999.
XX
PD
PF 30-APR-1999; 99WO-US009346.
XX
XX 01-MAY-1998; 98US-0083758P.
PR
PR 31-JUL-1998; 98US-0094869P.
```

```
PR 02-SEP-1998; 98US-0098994P.
PR 02-SEP-1998; 98US-0099062P.
PR 09-OCT-1998; 98US-0103749P.
PR 09-OCT-1998; 98US-0103794P.
PR 09-OCT-1998; 98US-0103796P.
PR 25-FEB-1999; 99US-0121528P.
XX
XX (CHIR ) CHIRON CORP.
PA (GENO-) INST GENOMIC RES.
XX
XX Fraser C, Galeotti C, Grandi G, Hickey E, Masignani V, Mora M;
PI Petersen J, Pizza M, Rappuoli R, Ratti G, Scalato E, Scarselli M;
PI Tettelin H, Venter JC;
XX
XX WPI; 2000-062150/05.
DR N-PSDB; AAZ54514.
XX
XX Novel Neisserial polypeptides predicted to be useful antigens for
XX vaccines and diagnostics.
XX
XX Claim 2; Page 1395; 1453pp; English.
XX
XX AAZ53015 to AAZ54536, AAZ54577 to AAZ54615, and AAY74253 to AAY75941
XX represent novel Neisseria meningitis and N. gonorrhoeae polynucleotides
XX and polypeptides. AAZ54537 to AAZ54576 and AAZ54616 to AAZ5473 represent
XX PCR primers used in the exemplification of the present invention. The
XX polypeptides, the polynucleotides, antibodies and compositions of the
XX invention can be used as vaccines, as diagnostic reagents, and as
XX immunogenic compositions. The polypeptides can be used in the manufacture
XX of medicaments for treating or preventing infection due to Neisserial
XX bacteria (e.g. meningitis and septicemia), to detect the presence of
XX Neisseria bacteria, or to raise antibodies. They may also be used to
XX screen for agonists or antagonists, which may themselves have use as
XX antibacterial agents. The polynucleotides of the invention may also be
XX used in gene therapy protocols
XX
SQ Sequence 508 AA;

Query Match          94.6%; Score 2573; DB 3; Length 508;
Best Local Similarity 97.8%; Pred. No. 5.5e-227;
Matches 497; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

QY 18 MKTRSLSLLCLLCCSSWLPPLLEERTESRHFNTPKPVLLDNLILQIRHTPHNGLSDIY 77
DB 1 MKTRSLSLLCLLCCSSWLPPLLEERTESRHFNTPKPVLLDNLILQIRHTPHNGLSDIY 60

QY 78 LLDDPHEAALARAALIESAHSLSLQYYIWRNDISGRLLFNMLYLAARGVVRVLLDDN 137
DB 61 LLNDPHEAFAARAALIESAHSLSLQYYIWRNDISGRLLFNMLYLAARGVVRVLLDDN 120

QY 138 NTRGLDDLLLALDSDHPNIEVRLFNPFVLRKWRALGYLTFPRLNRRMNKSFADNRATI 197
DB 121 NTRGLDDLLLALDSDHPNIEVRLFNPFVLRKWRALGYLTFPRLNRRMNKSFADNRATI 180

QY 198 LGGRNIGDEYFKVGEDTVFADLDIATGVSVEVSHDFDRYWASHAHNATRIIRSGNIG 257
DB 181 LGGRNIGDEYFKVGEDTVFADLDIATGVSVEVSHDFDRYWASHAHNATRIIRSGNIG 240

QY 258 KGLQALGYNDTSRHALLRYRETVEQSPLYQKIQTGRIDWQSVQRLISDTPAKGLDRDR 317
DB 241 KGLQALGYNDTSRHALLRYRETVEQSPLYQKIQTGRIDWQSVQRLISDTPAKGLDRDR 300

QY 318 RKPPIAGRLQDALKQPEKSVLVSPYFVPTKSGTDALAKLVQDGDIVTVLNSLOATDVA 377
DB 301 RKPPIAGRLQDALKQPEKSVLVSPYFVPTKSGTDALAKLVQDGDIVTVLNSLOATDVA 360

QY 378 AVHSGYVKYRKPPLKAGIKLYELOPNHAVPATKDKGLTGSSVTSLHAKTFIVDGKRIFIG 437
DB 361 AVHSGYVKYRKPPLKAGIKLYELOPNHAVPATKDKGLTGSSVTSLHAKTFIVDGKRIFIG 420

QY 438 SFNLDPRLSARLNTMGVVIESPKEIAEQMERTLADTTPPEYAVRVTLTKENRLQWHPDPAK 497
DB 421 SFNLDPRLSARLNTMGVVIESPKEIAEQMERTLADTTPPEYAVRVTLTKENRLQWHPDPAK 480
```



```
QY 498 TYPNPEAKLWKRIAAKILSLPIEGLL 525
Db 481 TYPNPEAKLWKRIAAKILSLPIEGLL 508

RESULT 6
ADL04719
ID ADL04719 standard; protein; 543 AA.
AC ADL04719;
XX
DT 06-MAY-2004 (first entry)
DE M. catarrhalis protein #485.
XX
KW Moraxella catarrhalis; infection.
XX
OS Moraxella catarrhalis.
XX
PN US6673910-B1.
XX
PD 06-JAN-2004.
XX
PF 04-APR-2000; 2000US-00540236.
XX
PR 08-APR-1999; 99US-0128416P.
XX
PA (GENO-) GENOME THERAPEUTICS CORP.
XX
PI Breton GL;
XX
DR WPI; 2004-178127/17.
DR N-PSDB; ADL02799.
XX
XX New nucleic acid encoding a Moraxella catarrhalis polypeptide, useful for
PT preparing a composition for diagnosing, preventing or treating infection
PT caused by Moraxella catarrhalis.
XX
PS Disclosure; SEQ ID NO 2405; 429pp; English.
XX
CC The invention relates to an isolated nucleic acid encoding an Moraxella
CC catarrhalis polypeptide. The nucleic acid is useful for preparing a
CC composition for diagnosing, preventing or treating infection caused by
CC Moraxella catarrhalis. The present sequence represents the amino acid
CC sequence of a M. catarrhalis protein.
XX
SQ Sequence 543 AA;

Query Match 30.6%; Score 831.5; DB 8; Length 543;
Best Local Similarity 35.9%; Pred. No. 6.3e-67;
Matches 194; Conservative 102; Mismatches 198; Indels 47; Gaps 12;

QY 22 SLISLLCLLLSCSSWLP-----PLEERTSRHNTSKPVLNLL-----QIRHT 67
Db 13 SIHTVMLVLLVGCKN-LPDTPLPKSLALTRQVQERHQMVKHGGSDGLISAINAQSHH 71
QY 68 PHNNGLSDIYLLDDPHEALAAARALIESAHSLLDQYIWRNDISGRLLFNLMYLAABRG 127
Db 72 PNQSG---YYPITGANAFARSTLTVDVANOSIDIQYIWHNDAGQLMLKDLWEAADRG 128
QY 128 VVRVRLLLDD-NNTRGDLLLALDSHPNIEVRFLENPFVLRKWRALGYLTDPPRLNRRMHN 186
Db 129 VIVRLLLDDFNSPELDQLLLRISKKNIAVRLINPMPYRGFRSLNMYLHPITNRRMHN 188
QY 187 KSFADNRPATILGRNIGDEYFKVGEDTVFADLILATGVSVGVSVDHFDRIYASHSAHN 246
Db 189 KSMTFDNKISVIGRNIIGNELNPNVHFNADLVMLVGHVVGKITQSFEIYASPLSFD 248
QY 247 ATRIIRSGNIG--KGLQALGYND-----ETSRHALLRYEETVEQSPLYQKIQTGIDW 297
Db 249 IETLVKHNKDDISGIKIPVWFDELEKVKDSSDADRELRTYRQAMQNSTIGDQLLAQVPF 308
```

```
QY 298 QSVOTRLISDTPAKGLDRDRKPPDIAGRLQDALKQPEKSVYLVSPYFVPTKSGTDAKL 357
Db 309 FWTIDLIADNVIKLNGHSDNEFLVSOLOQTOLGQPSKKUSIISYFVPTREGIDTLTL 368
QY 358 VQDGIDVTLTNSLOQATDVAHVSGYVKYRKPLKAGIKLYELOFN-HAVPATKDK--GL 414
Db 369 AKMGVEVRILTNSPDATDVGIVHAGYAHWRKQLLAAGVHLFEIKSSAQSIQDNENRFRWT 428
QY 415 TGSSTVSLHAKTFIVDGKRIPIGFSNLDPPSRARLNTMGVIVESPKIAEQHERTIADT-- 472
Db 429 RQHSTTSLHAKAFVDDDDQIFIGSYNDPPRSANINTELGVLIKDSKLAGQLHKALESQA 488
QY 473 TPEYAYRVTLDKHNRLOQHDHPATRKTPN-----EPEAKLWKRIAAKILSLPIEGL 524
Db 489 ITHQAYELKLDKAGNINWH-----TIENGQPVILHHEPHNMVHKDRIIWLGLMPDWL 542
QY 525 L 525
Db 543 L 543

RESULT 7
ABO75163
ID ABO75163 standard; protein; 564 AA.
XX
AC ABO75163;
XX
DT 29-JUL-2004 (first entry)
XX
DE Pseudomonas aeruginosa polypeptide #7338.
XX
KW Bacterial infection; Pseudomonas aeruginosa infection; antibacterial.
XX
OS Pseudomonas aeruginosa.
XX
PN US6551795-B1.
XX
PD 22-APR-2003.
XX
PF 18-FEB-1999; 99US-00252991.
XX
PR 18-FEB-1998; 98US-0074788P.
PR 27-JUL-1998; 98US-0094190P.
XX
PA (GENO-) GENOME THERAPEUTICS CORP.
XX
PI Rubenfield MJ, Nolling J, Deloughery C, Bush D;
XX
DR WPI; 2003-615309/58.
DR N-PSDB; ABD08734.
XX
PT Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide,
PT useful as molecular targets for diagnostics, prophylaxis and treatment of
XX pathological conditions resulting from bacterial infection.
PS Disclosure; SEQ ID NO 23909; 455pp; English.
XX
CC The invention relates to Pseudomonas aeruginosa polypeptides and the
CC polynucleotides encoding them. The sequences are useful in diagnosis and
CC therapy of pathological conditions, as molecular targets for diagnostics,
CC prophylaxis and treatment of pathological conditions resulting from a
CC bacterial infection, for evaluating a compound, such as a polypeptide,
CC for the ability to bind a P. aeruginosa nucleic acid, as components of
CC effective antibacterial targets, as targets for antibacterial drugs,
CC including anti-P. aeruginosa drugs, as templates for recombinant
CC production of P. aeruginosa-derived peptides or polypeptides, as target
CC components for diagnosis and/or treatment of P. aeruginosa-caused
CC infection, and in detection of P. aeruginosa sequences or other sequences
CC of Pseudomonas species using biochip technology. Sequences ABO67826-
CC ABO84396 represent P. aeruginosa polypeptides of the invention. Note: The
CC sequence data for this patent did not form part of the printed
CC specification but was obtained in electronic format from USPTO at
CC seqdata.uspto.gov/sequence.html
```



```

Oy 510 RIAAKILSLLP 520
Db 488 RLKEWSKVLP 498

RESULT 9
AAW55550
ID AAW55550 standard; protein; 502 AA.
XX
AC AAW55550;
XX
DT 24-JUN-1998 (first entry)
XX
DE H. pylori ORF 01ce21104_33203250_c3_87 secreted protein.
XX
KW Cytoplasmic; vaccine; prevention; treatment; infection; envelope;
KW identification; binding compound; bacteria; life cycle; activator;
KW inhibitor; duodenal ulcer disease; chronic gastritis; diagnosis.
XX
OS Helicobacter pylori.
XX
PN WO9737044-A1.
XX
FD 09-OCT-1997.
XX
PF 27-MAR-1997; 97WO-US005223.
XX
PR 29-MAR-1996; 96US-00625811.
PR 02-APR-1996; 96US-00758731.
PR 25-OCT-1996; 96US-00736905.
PR 28-OCT-1996; 96US-00738859.
PR 06-DEC-1996; 96US-00761318.
XX
PA (ASTR ) ASTRA AB.
XX
PI Smith D, Alm RA;
XX
DR WPI; 1997-503122/46.
DR N-PSDB; AAV24959.
XX
PT Helicobacter pylori nucleic acid sequences and encoded polypeptide(s) -
PT useful in vaccines to treat or prevent H. pylori infection and for
PT diagnosis of H. pylori infection.
XX
PS Claim 14,94; Page 755-756; 1145pp; English.
XX
CC This sequence is a H. pylori secreted protein. The protein may be used in
CC a vaccine to prevent or treat H. pylori infection or to identify H.
CC pylori polypeptide binding compounds, useful as potential H. pylori life
CC cycle activators or inhibitors. The DNA and probes derived from it may be
CC used for the identification of H. pylori in a sample and the diagnosis of
CC H. pylori infection. Nucleic acid sequences complementary to the DNA act
CC as antisense sequences and can be used to prevent the translation of H.
CC pylori mRNA. Antibodies against the protein can be used in immunoassays
CC to evaluate the abundance and distribution of H. pylori-specific
CC antigens. The genomic sequence of H. pylori (ATCC 55679) was determined
CC from overlapping contigs generated by mechanically shearing the bacterial
CC DNA. The sequences were analysed for ORF of at least 180 nucleotides, and
CC the predicted coding regions defined by computer evaluation. To identify
CC likely H. pylori antigens for vaccine development, the amino acid
CC sequences predicted from various ORF were analysed for significant
CC homology to other known or exported membrane proteins. Having identified
CC and determined the sequences of interest, particular regions can be
CC isolated from H. pylori by PCR amplification for recombinant polypeptide
CC production, e.g. in E. coli hosts
XX
SQ Sequence 502 AA;

Query Match 23.2%; Score 632; DB 2; Length 502;
Best Local Similarity 32.8%; Pred. No. 1.2e-48;
Matches 161; Conservative 93; Mismatches 181; Indels 56; Gaps 17;

64 IRHTPHNGLSDIY-----LLDDPHEALAAALATESAHSLSLDQYIWRN 109
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
30 ISYDPYTTTIGSLYAKNLKENPKHSAAILLEDGFDALLHRVGLIRMSQKSIDMTYIYKN 89
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
110 DISGRLLFNLMYLAERGVRRLLLDDNTRGLD-----DLLLALSDHPNIEVRLFNPPVL 165
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
90 DLSQVIAKELLNAANRGVVRILLDDN---GLSDSDFSDIML-LNPHKNIEVKIFNPYYI 145
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
166 RKWRALGY--LTDFFPLNRMRNKSFTADNRATILGRNIGTDEYFKVGGEDTVFADLDIL 222
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
146 RN-KGLRYFEMLADYERIKRMRNKLFIVDNFAVIIGRNIGDNYFNDLDTNFDLDDAL 204
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
223 ATGSVVGVEVSHDFDRYWASHSAHNATRIIRS-----GNIGKGLQALGYNDETSRHAL 274
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
205 FFGVASKAKESFENYWRFRHSIPVS-LLRTHKRLKNNVKEIAKLHEKIPISADANEFE 263
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
275 LRVRETVEQSPLYQ-KIQTRIDWQSVQTRLISDTPAKGLDRDRRKPPPIAGRLQDALQKP 333
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
264 KKVNDFTIERFQKYQYPIYG-----NAIFLADLPK-ITDPLYS-PIKIAFEKALKNA 314
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
334 EKSIVLVSPYFVPTKSCDATALKLVQDIDVTVTNSLQATDVAHVSHGVYKVRKPLKA 393
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
315 KDSVFTASSYFIPGKIMKIFKNOISKGIELNLTNSLSSTDAIVVYGAWERYENKLVRM 374
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
394 GIKLYELQPNHAVPATKDKGLTGSSVTSLSHAKTFIVDGKRFIFGSFNLDPRSARLNTMG 453
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
375 GANVYEIRNDFNRIQIGR---FSTKHSLSHGKTIIVFDDALTILGFSNIDPRSAYINTESA 431
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
454 VTIESPKIAEQMERTADTTPEYAYRVTLDKHNRLOMHDPATRKTY----PNEPEAKLWK 509
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
432 VLFNPNPSFAKRVRLSLKSHA-QQSMHLVLYRH-RVIWE--ATEBEGILHEKNSPDTSFEL 487
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
510 RIAAKILSLLP 520
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
488 RLKEWSKVLP 498

RESULT 10
AAW98378
ID AAW98378 standard; protein; 428 AA.
XX
AC AAW98378;
XX
DT 31-MAR-1999 (first entry)
XX
DE H. pylori GHPO 1375 protein.
XX
KW GHPO protein; Helicobacter infection; gastroduodenal disease; gastritis;
KW peptic ulcer disease.
XX
OS Helicobacter pylori.
XX
PN WO9843478-A1.
XX
PD 08-OCT-1998.
XX
PF 01-APR-1998; 98WO-US006371.
XX
PR 01-APR-1997; 97US-00833457.
PR 24-JUN-1997; 97US-00881227.
PR 29-JUL-1997; 97US-00902615.
XX
XX (INMR ) MERIEUX ORAVAX PASTEUR MERIEUX SERUMS.
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Kleanthous H, Al-Garawi A, Miller C, Tomb J, Oomen RP;
XX WPI; 1998-542293/46.
XX DR N-PSDB; AAX14037.
XX
XX New isolated Helicobacter polymucleotides - used to develop products for
XX the diagnosis, prevention and treatment of Helicobacter infections and
XX gastrointestinal diseases.

```


XX WPI; 2003-576092/54.
 DR N-PSDB; ADA30178.
 XX
 PT New Acinetobacter baumannii proteins and nucleic acids, useful as reagents
 PT for diagnosing a bacterial disease, as components of antibacterial
 PT vaccines, as targets for antibacterial drugs, or as biocontrol agents for
 PT plants.
 XX
 PS Example; SEQ ID NO 5591; 328pp; English.
 XX
 CC The invention relates to isolated Acinetobacter baumannii nucleic acids.
 CC The A. baumannii nucleic acids and polypeptides are useful as reagents
 CC for diagnosing a bacterial disease, as components of antibacterial
 CC vaccines, as targets for antibacterial drugs, to detect the presence of
 CC A. baumannii and other Acinetobacter species in a sample, in screening
 CC compounds for the ability to interfere with the A. baumannii life cycle
 CC or to inhibit A. baumannii infection, and as biocontrol agents for
 CC plants. The present sequence represents the amino acid sequence of an A.
 CC baumannii protein.
 XX
 SQ Sequence 275 AA;

Query Match 17.6%; Score 478.5; DB 6; Length 275;
 Best Local Similarity 41.0%; Pred. No. 5.9e-35;
 Matches 100; Conservative 51; Mismatches 78; Indels 15; Gaps 5;
 QY 294 RIDQSVQTRLIISDTPAKGLDRDRKPPKPIAGRLQDALKQKSVLYLSPYFVPTKSGTDA 353
 DB 35 KFDW--VIAEVVKDPSKIRSKAKKEHLNLFQNLHLEKPSNVLDLSAYFIPEKQAKI 92
 QY 354 LAKLVQDIDVTVLNTSLQATDVAHVHGVYKVRKLLKAGIKLYELQP-----NH 404
 DB 93 LSTLAKEGVEVRVLTNSFKANDVAVHAFYCKYKELLKNGVQLYEFLTPDKEDLNKNT 152
 QY 405 AVPATKQ-GLTGSSVTSLSHAKTFIVDGKRIFIGSFNLDPRSARLNTMGVWIESPKIAE 463
 DB 153 DELATKAKVNMKGLSRSLHTKLMALD-EQVFIFGSFNFDRPSAYLNTGIVLSDPSLAK 211
 QY 464 QMERTLADTTPYAYRVTLQKHNLQWHD--PATRKTYPNEPEAKLWKRTAAKILSLPLI 521
 DB 212 TIHTMDENLNKYAKLKLQPNNHIIYWOQETPKGVYKKEPKMKWKQAGMKLLSLWPL 271
 QY 522 EGLL 525
 DB 272 EGFN 275

RESULT 13
 ID ABM88218 standard; protein; 2519 AA.
 AC ABM88218;
 XX
 DT 02-JUN-2005 (first entry)
 XX
 DE Rice abiotic stress responsive polypeptide SEQ ID NO:6464.
 XX
 KW abiotic stress tolerance; transgenic plant; plant; cereal; agriculture.
 XX
 OS Oryza sativa.
 XX
 PN WO2003008540-A2.
 XX
 PD 30-JAN-2003.
 XX
 PF 21-JUN-2002; 2002WO-US019668.
 XX
 PR 22-JUN-2001; 2001US-0300112P.
 PR 24-AUG-2001; 2001US-0314662P.
 PR 26-SEP-2001; 2001US-0325277P.
 PR 21-NOV-2001; 2001US-0332132P.
 XX

PA (SYGN) SYNGENTA PARTICIPATIONS AG.
 XX
 PI Kreps J, Briggs SP, Cooper B, Glazebrook J, Goff SA, Katagiri F;
 PI Moughamer T, Provart N, Ricke D, Zhu T;
 XX
 XX WPI; 2003-248011/24.
 XX
 PT New stress-responsive nucleic acid, useful for altering the
 PT responsiveness of a plant, e.g. cereal, to an abiotic stress such as cold
 PT stress, salt stress or osmotic stress.
 XX
 PS Claim 1; SEQ ID NO 6464; 89pp; English.
 XX
 CC The invention relates to novel abiotic stress responsive polynucleotides
 CC and polypeptides. Also disclosed are vectors, expression cassettes, host
 CC cells, and plants containing such polynucleotides. Also disclosed are
 CC methods for using the polynucleotides and polypeptides to alter the
 CC responsiveness of a plant to abiotic stress. The invention is useful in
 CC agriculture. The nucleic acid is useful for determining whether a test
 CC plant has been exposed to an abiotic stress condition. It is also useful
 CC for selecting an agent that alters abiotic stress regulated
 CC polynucleotide expression in a plant cell, and to identify a homolog or
 CC ortholog to an abiotic stress responsive polynucleotide. The nucleic acid
 CC molecule and the polypeptide encoded by it are useful in altering the
 CC responsiveness of a plant to an abiotic stress, such as cold stress, salt
 CC stress, osmotic stress or any of their combinations. The present sequence
 CC is used in the exemplification of the invention
 XX

SQ Sequence 2519 AA;

Query Match 15.4%; Score 419; DB 7; Length 2519;
 Best Local Similarity 28.2%; Pred. No. 5.5e-28;
 Matches 147; Conservative 76; Mismatches 186; Indels 112; Gaps 18;
 QY 32 CSCSSLPPLLEER--TES-----RHFNTPKPVLLDNILQIRHTPHNGLSDIYLLDDPHE 84
 DB 377 CAQDSPLRALAGRAFTSTPEQPRHYAT-----LLDEGEG 411
 QY 85 ALAARAALTESAEHSLDLQYIWRNDISGRLLFNLMYLAERGVVRVLLDDNNTRGLDD 144
 DB 412 ALVARLNLRSATRSIDLQTYIFDKDSDARLVDELAAASRRGVKRVLDQLSAISDLQ 471
 QY 145 LLLALD-SHPNIEVRILFNP-FVLKRWALGY----LTDFFLNRRHMKSFSTADNRATIL 198
 DB 472 ILGALSGAQNFQLRVYNPTFGKARLNYFDYAGSVLCCCFRFRNORMHKNLLVDDAIGVV 531
 QY 199 GGRNIGDEYFKVGEDTVFADLDILATGSSVVGEVSHDFDRIYASHASHAHNATRIIRSGNIGK 258
 DB 532 GGRNYQDDYYDWDREYNFRDRDLVIAGPEARAMAANFADFHWARRSVPAERL---NDVGR 588
 QY 259 GLQALGVNDETSRHALRYRETYEQSPLYQKIQTGRIDWQS-----VQTRLSIDT 308
 DB 589 TLLREGV--PTLPPASFRFRPVROR----VSAENDMDVFSRFSVDTALPVASVRYVADL 642
 QY 309 PAKGLDRDRRKPPPIAG-----RLQDALKQKPSVYLVSPYFVPTKSGTDALAKLVQ--D 360
 DB 643 PRKH-RREKADAPLAGQHVTEPQOLDALIAGAQEVILQTPYLVLSPAQKLFRELKRPQ 701
 QY 361 GIDVTVLNLSLQATDVAHV-----SGVYKYPKLLKAG 394
 DB 702 PPRVVVSSNSLAATDNPVYALSYNDRDRDEHPAPTQKQLAVNPGTGSAYQNAERERRAA 761
 QY 395 IKL-----YELQP-----NHAVPATKDKLGTSSVTSLSHAKTFIVDGKRIFTGSPN 440
 DB 762 SEVETLLRTETRPSFLGSAVKNKLPVTRKGRMG-----LHAKSLWDRRIIGVVGTHN 816
 QY 441 LDPRSARLNTMGVWIESPKIAEQM-ERTLADTTPYAYRV 480
 DB 817 FDRSENYNTEGAVIIDDPAFAEQLAESIIRDHPQNSWTV 857

RESULT 14
 ADA36398

```
ID ADA36398 standard; protein; 321 AA.
XX
AC ADA36398;
XX
DT 20-NOV-2003 (first entry)
XX
DE Acinetobacter baumannii protein #3559.
XX
KW Acinetobacter baumannii; bacterial disease; antibacterial; vaccine;
KW plant biocontrol agent.
XX
OS Acinetobacter baumannii.
XX
PN US562958-B1.
XX
PD 13-MAY-2003.
XX
PF 04-JUN-1999; 99US-00328352.
XX
PR 09-JUN-1998; 98US-0089701P.
XX
PA (GENO-) GENOME THERAPEUTICS CORP.
XX
PI Breton G, Bush D;
XX
DR WPI; 2003-576092/54.
XX
DR N-PSDB; ADA32272.
XX
New Acinetobacter baumannii proteins and nucleic acids, useful as reagents
PT for diagnosing a bacterial disease, as components of antibacterial
PT vaccines, as targets for antibacterial drugs, or as biocontrol agents for
PT plants.
XX
PS Example; SEQ ID NO 7685; 328pp; English.
XX
The invention relates to isolated Acinetobacter baumannii nucleic acids.
CC The A. baumannii nucleic acids and polypeptides are useful as reagents
CC for diagnosing a bacterial disease, as components of antibacterial
CC vaccines, as targets for antibacterial drugs, to detect the presence of
CC A. baumannii and other Acinetobacter species in a sample, in screening
CC compounds for the ability to interfere with the A. baumannii life cycle
CC or to inhibit A. baumannii infection, and as biocontrol agents for
CC plants. The present sequence represents the amino acid sequence of an A.
CC baumannii protein.
XX
SQ Sequence 321 AA;
Query Match 15.0%; Score 409; DB 6; Length 321;
Best Local Similarity 34.7%; Pred. No. 1.8e-28;
Matches 100; Conservative 44; Mismatches 110; Indels 34; Gaps 6;
QY 30 LLCSSWL--PPLERTESRHNTSKPVLNDNLIQIRTHPHNGLSDIYLLDDPHEALA 87
D: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
D: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
44 IVIKTNWLNDEAAEKTQ-----QGLTAFVALDDAFMSIA 79
QY 88 ARAALIESAHSLDILOYIWRNDISGRLLFNLMYLAARGVYRVLILDDNTRGLDILL 147
D: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
D: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
80 SRIYLIRNAKETIDLOYYIWTNDFVGNLTHLHLLKAADRGIKVRLIDDQNGIKLDGILR 139
QY 148 ALDSDHPNIEVLFNPFVLKRWALGYLTFPRLNRRMKNKSFADNRATILGGRNIGDEY 207
D: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
D: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
140 SLLOHTNFEIRLFNPKYRFLRIFDYLPFRKVNHRMKNKLIADASTVGTGRNISSEY 199
QY 208 FKVGEDTVFADLILATGSVGEVSHDFDRYWASHSAHNATRIIRSGNIGK-GLQALGVN 266
D: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
D: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
200 FEASSKFQTDMDILFYGHAVRHAQAVFTDFWESTLSVNAIEII--GTCAEHLKALREH 257
QY 267 DET---SRHLLRYRETVEQSPLYQKIQTGRIDWQSVQTRILSDTPAK 311
D: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
D: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
258 YEQLHHEDHSUTEDKLYDAQSYLKELLEHNPIQWS--KAHFVADSPKK 303
RESULT 15
```

```
ABM71816
ID ABM71816 standard; protein; 513 AA.
XX
AC ABM71816;
XX
DT 20-NOV-2003 (first entry)
XX
DE Staphylococcus aureus protein #1056.
XX
KW Antibacterial; vaccine; gene therapy; infection; sepsis; diagnosis;
KW enzymatic assay; antibiotic target.
XX
OS Staphylococcus aureus.
XX
PN WO200294868-A2.
XX
PD 28-NOV-2002.
XX
PF 27-MAR-2002; 2002WO-IB002637.
XX
PR 27-MAR-2001; 2001GB-00007661.
XX
PA (CHIR-) CHIRON SPA.
XX
PI Masignani V, Mora M, Scarselli M;
XX
DR WPI; 2003-120786/11.
XX
DR N-PSDB; ACF73376.
XX
New Staphylococcus aureus protein, useful as a vaccine for treating or
PT preventing Staphylococcal infection, specifically an infection caused by
PT S. aureus, e.g. sepsis.
XX
PS Claim 1; SEQ ID NO 2112; 49pp; English.
XX
The invention relates to novel genes and encoded proteins from
CC Staphylococcus aureus. A composition comprising the S. aureus protein, a
CC nucleic acid encoding the protein, or an antibody to the protein, is
CC useful as a pharmaceutical, particularly as a vaccine for treating or
CC preventing infection due to Staphylococcus bacteria, specifically an
CC infection caused by S. aureus. The composition is particularly useful for
CC treating or preventing sepsis in a patient. The composition can also be
CC used for diagnostics. The protein is also used in an assay for enzymatic
CC studies and as a target for antibiotics. This sequence represents one of
CC the novel S. aureus proteins of the invention
XX
SQ Sequence 513 AA;
Query Match 11.3%; Score 308; DB 6; Length 513;
Best Local Similarity 23.2%; Pred. No. 7.2e-19;
Matches 119; Conservative 76; Mismatches 155; Indels 162; Gaps 19;
QY 4 DPKIQAMPSETISP-----MKTRSLISLLCULLCSCSWLPPLEBTERSHFNTSKPVLL 58
D: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
D: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
110 DEQLAALKNENFNSNYQIVKFKEMIQMLL-----YNNAAFLTT 148
QY 59 DNILQIRTHPHNGLSDIYLLDDPHEALARAALIESAHSLDILOYIWRNDISGRLLFN 118
D: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
D: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
149 DNDLKI-YTDGQEKFPD--LIQD-----IRNATDYTHFOYYIIONDELGRITLN 194
QY 119 LMYLAARGVYRVLILDDNTRGLDILLALDSDHPNIEVLFNPFVLKRWALGYLTDFF 178
D: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
D: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
195 ELGKAEQGVKLYIDDMGSRGL-----RKGLRPPRNGKHAEAF--FP 238
QY 179 -----RLNRRMHNKSFTADNRATILGGRNIGDEYFKVGEDTVFA---DLILATGSV 227
D: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
D: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
239 SKPLPLNLRMNNRHKIWIWDGQIGYGVGFNVGDEY--LGKSKKFGYWRDTHLRIVGDA 296
QY 228 VGEVSHDFDRYWASHSAHNATR-----IIRSGNIGKGLQALGVNDETSRHAL 275
D: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
D: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
297 VNALQLRFLDWN-----QATRDHISYDDRVFPDWNVSGGTIGVQIASSGPDDE----- 345
QY 276 RYRETVEQSPLYQKIQTGRIDWQSVQTRILSDTPAKGLDRDRRKPPPIAGRLQDALKQPEK 335
```

[illegible]

RESULT 16	
ADG30702	
ID	ADG30702 standard; protein; 486 AA.
XX	
AC	ADG30702;
XX	
DT	26-FEB-2004 (first entry)
XX	
DE	Xanthomonas axonopodis pv citri plant pathology-related XAC0014 protein.
XX	
KW	Xanthomonas microorganism; plant; pathology; bacterial pest; Xac; Xcc;
KW	XAC.
XX	
OS	Xanthomonas axonopodis pv. citri.
XX	
PN	WO2003089647-A1.
XX	
PD	30-OCT-2003.
XX	
PF	22-APR-2003; 2003WO-BR000060.
XX	
PR	22-APR-2002; 2002US-0374620P.
XX	
PA	(AMPA-) FUNDACAO AMPARO A PESQUISA DO ESTADO.
XX	
PI	Da Silva ACR, Farah SC, Quaggio RB, Reinach PDC, Ferro JA;
PI	De Oliveira JCF, De Laia ML, Setubal JC, Furian LR;
XX	
DR	WPI; 2003-865444/80.
DR	N-ESDB; ADG30701.
XX	
PT	New nucleic acid molecule from a Xanthomonas microorganism, useful in
PT	determining the presence of Xanthomonas bacteria in a sample.
XX	
PS	Claim 8; SEQ ID NO 59; 145pp; English.

PS Claim 8; SEQ ID NO 59; 145pp; English.

The invention relates to a novel isolated nucleic acid molecule from a Xanthomonas microorganism where the nucleic acid molecule is associated with pathogenicity caused by the Xanthomonas microorganism, or its variant, that causes reduced or enhanced pathogenicity. The nucleic acid of the invention may be useful in detecting the presence of Xanthomonas bacteria in a sample, as well as in plant pathology, for example, for identifying nucleic acid molecules and proteins involved in pathology caused by bacterial pests. The current sequence is that of the Xanthomonas axonopodis pv. citri (Xac) plant pathology-related XAC protein of the invention.

Sequence 486 AA;

Query Match 10.9%; Score 296; DB 7; Length 486;
Best Local Similarity 25.2%; Pred. NO. 8.4e-18;
Matches 123; Conservative 63; Mismatches 209; Indels 94; Gaps 14;

Qy	33	SCSSWLPLEERTESRHFNTSKPVLDDNILQIRHTPHNGLSDIYLLDDPHEALARAAL	92
		: :	
Db	89	SYSSVCPPDADCTE-----LAKIAOSTTGLAPSSATEVHVLVDGAATYAATIEA	137

QY	93	IESAEHSJDLQYIWRNDISGRLLFNMLYAAERGVRVRLLLDDNNTRGLDLLLALD	152
Db	138	IRGARDHITHLEYIIFOPDHSGTACAAALMERARAGVKVRLMLDAIGSSAMTRALRTRE	197
QY	153	PNIEVRLFNPFVLKWRALGYLTDF--PRLNPRMKNKSFADNRATILGGRNIGDEYFKV	210
Db	198	AGVETAWFHPSQL-----LKPKFRPWLNRTHRKVIIDGRIGFTGGIIVNTDDENEQ	249
QY	211	GEDTVFADLDILATGSSVVGVESHDFRYWASHSAHNATRIIRSGNIGKGLQALGYNDETS	270
Db	250	VRKQAYRDLHVRQLGVHVSRLQVLFEDWL-----YATSQGRAAFHGQOL-----	300
QY	271	RHALLRYBETVEQSPLYOKIQTGRIDWQSVQTRLISDTPAKGLDRDRKPPIAGRLQDAL	330
Db	301	TRA-----QGTVDQAVL-----VSGPDSSWEAIIHRLMVA-----AI	331
QY	331	KQEPSKVLVSPFYFPTKSGTDLAKLVQDGDIVTLTNSLQATDVAAVHSGVYKVRKPL	390
Db	332	HEAKHRVWLVTYPYFVPGEAARWALTSAALGGLDVRLVLP--RVSDRLVTYAARSYFDEL	389
QY	391	LKAGIKLYELOPNHIAVPATKDKGLTGSSVTSLSHAKTFIVDGKRIFIGSFNLDPRSARLNT	450
Db	390	LEAGVRIIYEYGR-----MLHTKALLADDVCIVGSANFDSFRFLNF	432
QY	451	EMGVVIESPKIAEQMERTLADTTPEYAYRVTLDKHNRLOWHDPATREKTYNPEEAKLWKR	510
Db	433	ELSMLFRDQVAAM--AGLIGTDIQQAOEQVFAHRPL--WRS-----RUPEA-----	477
QY	511	IAAKILSL	519
Db	478	-FARLLSPL	485

RESULT 17

ABB49813
ID ABB49813 standard; protein; 504 AA.

XX
AC

XX
DT 05-FEB-2002 (first entry)XX
DE Listeria monocytogenes protein #2517.

XX Antibacterial; gene therapy; vaccine; biosynthesis; biodegradation;
KW vitamin B12; bacterial infection; disease.
KW

AA
OS
Listeria monocytogenes.

XX PN WO200177335-A2.

XX
PD
18-OCT-2001XX
PF 11-APR-2001:XX
11-APR-2000. 2000EP-00004528
PPXX
PA (TNCB) TNCB PACETIM

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100	101	102	103	104	105	106	107	108	109	110	111	112	113	114	115	116	117	118	119	120	121	122	123	124	125	126	127	128	129	130	131	132	133	134	135	136	137	138	139	140	141	142	143	144	145	146	147	148	149	150	151	152	153	154	155	156	157	158	159	160	161	162	163	164	165	166	167	168	169	170	171	172	173	174	175	176	177	178	179	180	181	182	183	184	185	186	187	188	189	190	191	192	193	194	195	196	197	198	199	200	201	202	203	204	205	206	207	208	209	210	211	212	213	214	215	216	217	218	219	220	221	222	223	224	225	226	227	228	229	230	231	232	233	234	235	236	237	238	239	240	241	242	243	244	245	246	247	248	249	250	251	252	253	254	255	256	257	258	259	260	261	262	263	264	265	266	267	268	269	270	271	272	273	274	275	276	277	278	279	280	281	282	283	284	285	286	287	288	289	290	291	292	293	294	295	296	297	298	299	300	301	302	303	304	305	306	307	308	309	310	311	312	313	314	315	316	317	318	319	320	321	322	323	324	325	326	327	328	329	330	331	332	333	334	335	336	337	338	339	340	341	342	343	344	345	346	347	348	349	350	351	352	353	354	355	356	357	358	359	360	361	362	363	364	365	366	367	368	369	370	371	372	373	374	375	376	377	378	379	380	381	382	383	384	385	386	387	388	389	390	391	392	393	394	395	396	397	398	399	400	401	402	403	404	405	406	407	408	409	410	411	412	413	414	415	416	417	418	419	420	421	422	423	424	425	426	427	428	429	430	431	432	433	434	435	436	437	438	439	440	441	442	443	444	445	446	447	448	449	450	451	452	453	454	455	456	457	458	459	460	461	462	463	464	465	466
---	---	---	---	---	---	---	---	---	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----

PI Dussurget O, Chetoui F, Nedjari H, Glaser P, Kunst F, Cossart P;
PI Daniels J, Goebel W, Kraft J, Kuhn M, Ng E, Vazquez-Boland JA;
PI Dominguez-Bernal G, Garrido-Garcia P, Tierrez-Martinez A, Anend A;
PI Chakraborty T, Domann E, Hain T, Berche P, Charbit A, Durant L;
PI Perez-Diaz J, Baquero F, Garcia Del Portillo F, Gomez-Lopez N;
PI Madueno E, De Pablo B, Wehlund J, Kaerst U, Entian K, Hauf J;
PI Rose M, Voss H;

XX
DR WPI; 2002-010914/01.

XX	Genomic sequence for <i>Listeria</i> monocytogenes, useful e.g. for treatment
PT	PT and prevention of <i>Listeria</i> and related bacterial infections, and related
PT	polypeptides.


```

Db      399 EAGVNVVHY-----DNGF-----LHSEKTLVIDEASVGTANMDNRSFTLNFE 441
Qy      452 MGVIIESPKIAEQMERTLAD-----TPPEYAYRVTLDK 484
Db      442 VNAFIYDEGVARSLSQAFINDMKLSNKLTSSEYAKRNLLVK 482

RESULT 19
ADS04559
ID      ADS04559 standard; protein; 495 AA.
XX
AC      ADS04559;
XX
DT      04-NOV-2004 (first entry)
XX
DE      Staphylococcus epidermis polypeptide seqid 3854.
XX
KW      antibacterial; vaccine; antisense therapy; Staphylococcus epidermidis;
KW      recombinant expression vector; infection; computer readable medium;
KW      computer based system.
XX
OS      Staphylococcus epidermidis.
XX
PN      US2004147734-A1.
XX
PD      29-JUL-2004.
XX
PF      01-DEC-2003; 2003US-00724972.
XX
PR      08-NOV-1997; 97US-0064964P.
PR      13-AUG-1998; 98US-00134001.
PR      29-NOV-1999; 99US-00450969.
XX
PA      (DOUC/) DOUCETTE-STAMM L.
PA      (BUSH/) BUSH D.
XX
PI      Doucette-Stamm L, Bush D;
XX
DR      WPI; 2004-580138/56.
XX
N-PSDB; ADS00787.
XX
PT      New isolated polypeptide and encoding nucleic acid derived from
PT      Staphylococcus epidermidis, useful for diagnosing, preventing and/or
PT      treating an S. epidermidis bacterial infection.
XX
PS      Claim 17; SEQ ID NO 3854; 741pp; English.
XX
CC      The invention describes an isolated nucleic acid comprising a nucleotide
CC      sequence with any of 3772 fully defined nucleotide sequences (SEQ ID NO:
CC      1-3772) and encoding an Staphylococcus epidermidis polypeptide with any
CC      of 3772 fully defined amino acid sequences (SEQ ID NO: 3772-7544) as
CC      given in the specification. Also described are: a recombinant expression
CC      vector; a cell comprising a recombinant expression vector of (1);
CC      producing an S. epidermidis polypeptide; an isolated nucleic acid
CC      comprising a nucleotide sequence of at least 8 nucleotides in length; a
CC      vaccine composition for prevention or treatment of an S. epidermidis
CC      infection, comprising a nucleic acid cited above and a carrier; treating
CC      a subject for S. epidermidis infection; a recombinant or substantially
CC      pure preparation of an S. epidermidis polypeptide or its fragment; a
CC      vaccine composition for prevention or treatment of an S. epidermidis
CC      infection; detecting the presence of a Staphylococcus nucleic acid in a
CC      sample; a computer readable medium having recorded in it the nucleotide
CC      sequences with SEQ ID NO: 1-3772 or its fragments; a computer based
CC      system for identifying fragments of the Staphylococcus genome of
CC      commercial importance; a computer based system for identifying fragments
CC      of the Staphylococcus plasmids of commercial importance; identifying
CC      commercially important nucleic acid fragments of the Staphylococcus
CC      genome and/or plasmids; and identifying an expression modulating fragment
CC      of the Staphylococcus genome and/or plasmids. The methods and
CC      compositions of the present invention are useful for the diagnosis,
CC      prevention and/or treatment of an Staphylococcus epidermidis bacterial
CC      infection. This is the amino acid sequence of a S. epidermis protein of

```

```

CC      the invention.
XX
SQ      Sequence 495 AA;
XX
Query Match      10.7%; Score 291; DB 8; Length 495;
Best Local Similarity 21.5%; Pred. No. 2.5e-17;
Matches 112; Conservative 79; Mismatches 166; Indels 164; Gaps 17;

Qy      1 MHTDPKIQAMPSETISP-----MKTRSLISLILCLLSCSSWLPPLERTESRFTNTPSKP 55
Db      89 MIVDEQLEALKQDFSGKNHQIVKFKEMVQMLL-----YNNAAF 127
Qy      56 VLLDNILQIRHTPHNGLSDIYLLDDPHEALAAALIESAEHSLDLYIWRNDISGRLL 115
Db      128 LTTDN-----DLTIYDGHQKFDLINDIRHAQSVIHYIYIIHSDNLGKQ 173
Qy      116 LFNLMYLAERGVRVRLLLDDNNTRGLDLLLLALDHPNIEVRLFNPFVLRKWRALGYLT 175
Db      174 LLHELEKKAEGIEVKWLYDDMGSRDL-----RKDKLKKFRQKGGA 215
Qy      176 D--FP-----RLNRRMHKSFADNRATILGGRNIGDEYFKVGEDTVFA---DLIDL 222
Db      216 ESFFPSKLPILNLRMMNRNRKIVIDGTTGYVGGFNVGDEY--IGSKKFGYWRDTHLR 273
Qy      223 ATGSVVGEVSHDFRYWASHSAH-NATRIIR-----SGNIGKGLQALGYNDSTSRHAL 274
Db      274 IKGDVNALQLRFLDWNQSQTRDNLTYESYFPDVSDDGTTIGIQAISGPDE----- 326
Qy      275 LRYRETVESPLYOKIQTGRIDWQSVQ---TRUISDTPAKGLDRDRRKPPIAGRLQALK 331
Db      327 -----DWEQIKYGLKMISSA----- 342
Qy      332 QPEKSVLYSPYEVPTKSGTDLAKLVQDGDIVTLTNSLQATDVAHVHSGYVKYKPELL 391
Db      343 --KESIYQSPYFIPDQAFLDISKIALGQVDVNMVFNKR--DHPFVYVWATLKNVASLL 398
Qy      392 KAGIKLYELOPNHAPVATPKDKGLTSGSVTSLHAKTFIVDGKIFIGSFNDPRSARLNTE 451
Db      399 EAGVNVVHY-----DNGF-----LHSEKTLVIDEASVGTANMDNRSFTLNFE 441
Qy      452 MGVIIESPKIAEQMERTLAD-----TPPEYAYRVTLDK 484
Db      442 VNAFIYDEGVARSLSQAFINDMKLSNKLTSSEYAKRNLLVK 482

RESULT 20
ABB49993
ID      ABB49993 standard; protein; 482 AA.
XX
AC      ABB49993;
XX
DT      05-FEB-2002 (first entry)
XX
DE      Listeria monocytogenes protein #2697.
XX
KW      Antibacterial; gene therapy; vaccine; biosynthesis; biodegradation;
KW      vitamin B12; bacterial infection; disease.
XX
OS      Listeria monocytogenes.
XX
PN      WO200177335-A2.
XX
PD      18-OCT-2001.
XX
PF      11-APR-2001; 2001WO-FR001118.
XX
PR      11-APR-2000; 2000FR-00004629.
XX
PA      (INSP ) INST PASTEUR.
XX
PI      Buchrieser C, Frangeul L, Couve E, Rusniok C, Feihi H, Dehoux P;
PI      Dussurget O, Chetouani F, Nedjari H, Glaser P, Kunst F, Cossart P;
PI      Daniels J, Goebel W, Kreft J, Kuhn M, Ng E, Vazquez-Boland JA;

```


PS Example 2; SEQ ID NO 84; 174pp; English.

XX The invention relates to methods of identifying an agent that binds a polypeptide from *Staphylococcus aureus* by contacting a polypeptide and an agent to form a mixture and determining whether the agent binds the polypeptide. The method is useful in identifying agents that bind gene products critical for the survival of microbes, preferably *Staphylococcus* microbes, including agents that interfere with the function of such gene products. The method is used to identify new agents useful for treating or preventing bacterial infections, particularly those caused by antibiotic-resistant bacteria. This sequence represents a protein from *Staphylococcus aureus* of the invention when the corresponding gene is cloned into the expression vector pQ8-60 or pQ8-70. This encoding gene sequence corresponds to the essential gene shown in ADC21335.

SQ Sequence 502 AA;

Query Match 10.5%; Score 286.5; DB 7; Length 502;
Best Local Similarity 21.1%; Pred. No. 6.6e-17;
Matches 106; Conservative 86; Mismatches 169; Indels 141; Gaps 16;

QY 42 EERTSRHFTSKPVLLD-----NILQIRHTPHNGLSDIYLLDDPHEALAAARALIES 95
Db 105 DNKQVQKHDLVRMLMDQDGLTENNKKVDHFIDGNDLYDQVLKD-----IKN 152
QY 96 AEHSLDLQYIYWRNDISGRLLFNMLAAERGVRVRLLLDDNNTRGLDDLLALDHPNI 155
Db 153 AKEYIHLEYTYTFALDGLGKRILHALEEKLGKQLEVKILYDDVGSK-----NV 199
QY 156 EVRLFNPPVLRKWRALG-----YLTDFP-----RLNRRMHNKSFADNRATILGGRNIGD 205
Db 200 KMANFDHF-----KSLGGEVAFASKLPLNLFNRRNNRHKIIVIDQGLGVGGFNIGD 254
QY 206 EYFKVGEDTVFADLDILATGSGVGVSHDFDRYVWASHSAHNATRII-----RSGNIGK 258
Db 255 EYLGKGLGYWRDTHLRIGQDAVDALQRLFDLWNS-QAHRPQFEYDKVYFKKNGPLG- 312
QY 259 GLQALGYNDTSRHALLRYRETVEOSPLYQKIQTGRIDWQSVQ---TRLISDTPAKGLDR 315
Db 313 -----NSPTQIAASGPASDWHQIEGYTKMMSA----- 341
QY 316 DRKPPPIAGRLQDALKQPEKSVYLSVPVPTKSGTDALAKLVODIGDVTVLTSLQATD 375
Db 342 -----KKSIVLOSPIYFIDNSVINAIIKAAGSGVDVHLMIPC--KPD 381
QY 376 VAAHSGYVKYRKPILLKAGIKLYELQPNHAVPATPKDGLTSSVTSLSHAKTFIVDGKRF 435
Db 382 HPLVYVWATFSNASDLSSGVKIY-----TVENGFG-----IHSKMKLIDDEIVS 424
QY 436 IGSFNLDPRSAKLNTMGVVIESPKIARQMBERTLADTTPYAYRVTLDKHRLQWHDPAT 495
Db 425 VGTANMDFRSEFNLFNFAVYVDENLAKDL-----RVAYEHDIITKSKQL-----T 469
QY 496 RKTYPNEPEAKLWKRIAAKILS 517
Db 470 KESYANRPLSVKFESLAKILS 491

RESULT 27
ADC25088
ID ADC25088 standard; protein; 502 AA.
XX
AC ADC25088;
XX
DT 18-DEC-2003 (first entry)
XX
DE *S. aureus* polypeptide #9 encoded by the essential coding region.
XX
KW Binding agent; growth rate; microbe; virulence; vaccine;
KW bacterial infection; antibiotic-resistant bacteria; bacteraemia;
KW septic shock; metastatic infection; endocarditis; arthritis;
KW osteomyelitis; pneumonia; abscess; skin rash; food poisoning;
KW multisystem dysfunction; toxic shock syndrome; antimicrobial;

KW antibacterial; immunosuppressive; cytostatic; antiinflammatory.
XX
OS *Staphylococcus aureus*.
XX
FN US2003087321-A1.
XX
PD 08-MAY-2003.
XX
PF 28-SEP-2001; 2001US-00966521.
XX
PR 28-SEP-2001; 2001US-00966521.
XX
PA (TOMI/) TOMICH C C.
PA (QUIN/) QUINN C L.
PA (ARVI/) ARVIDSON S.
PA (MOTT/) MOTT J E.
PA (HARR/) HARRIS D W.
XX
PI Tomich CC, Quinn CL, Arvidson S, Mott JE, Harris DW;
XX
XX WPI; 2003-606127/57.
XX
DR N-PSDB; ADC25087.
XX
PT Identification of agent that binds polypeptide, for treating bacterial infections, particularly caused by antibiotic-resistant bacteria,
PT involves mixing polypeptide and agent, and determining whether agent binds polypeptide.
XX
PS Example 2; SEQ ID NO 84; 122pp; English.
XX
CC The invention discloses a method for identifying an agent that binds a polypeptide which comprises mixing a polypeptide and an agent to form a mixture and determining whether the agent binds the polypeptide. Also
CC claimed is a method for identifying an agent that decreases the growth rate of a microbe, making a *Staphylococcus aureus* with reduced virulence, which comprises altering a coding sequence in an *S. aureus* to introduce a mutation, a vaccine composition comprising the *S. aureus* organism with reduced virulence. The methods are useful for identifying an agent that binds a polypeptide, useful in treating (e.g. as a vaccine) bacterial infections, particularly those caused by antibiotic-resistant bacteria, including bacteraemia, septic shock and serious metastatic infections including endocarditis, arthritis, osteomyelitis, pneumonia, abscesses in virtually any organ, skin rashes, food poisoning or multisystem dysfunction, i.e. toxic shock syndrome. The inventive method identifies agents that are unrelated to existing antimicrobials and that target different aspects of *Staphylococcus aureus* invasion and replication in the host. The sequence presented is one of the *S. aureus* polypeptides of the invention encoded by the essential coding region.

SQ Sequence 502 AA;

Query Match 10.5%; Score 286.5; DB 7; Length 502;
Best Local Similarity 21.1%; Pred. No. 6.6e-17;
Matches 106; Conservative 86; Mismatches 169; Indels 141; Gaps 16;

QY 42 EERTSRHFTSKPVLLD-----NILQIRHTPHNGLSDIYLLDDPHEALAAARALIES 95
Db 105 DNKQVQKHDLVRMLMDQDGLTENNKKVDHFIDGNDLYDQVLKD-----IKN 152
QY 96 AEHSLDLQYIYWRNDISGRLLFNMLAAERGVRVRLLLDDNNTRGLDDLLALDHPNI 155
Db 153 AKEYIHLEYTYTFALDGLGKRILHALEEKLGKQLEVKILYDDVGSK-----NV 199
QY 156 EVRLFNPPVLRKWRALG-----YLTDFP-----RLNRRMHNKSFADNRATILGGRNIGD 205
Db 200 KMANFDHF-----KSLGGEVAFASKLPLNLFNRRNNRHKIIVIDQGLGVGGFNIGD 254
QY 206 EYFKVGEDTVFADLDILATGSGVGVSHDFDRYVWASHSAHNATRII-----RSGNIGK 258
Db 255 EYLGKGLGYWRDTHLRIGQDAVDALQRLFDLWNS-QAHRPQFEYDKVYFKKNGPLG- 312
QY 259 GLQALGYNDTSRHALLRYRETVEOSPLYQKIQTGRIDWQSVQ---TRLISDTPAKGLDR 315

DE M. xanthus protein sequence, seq id 16463.
KW Transgenic plant; DNA replication; gene regulation; gene expression.
OS Myxococcus xanthus.
XX US6833447-B1.
FN 21-DEC-2004.
XX 10-JUL-2001; 2001US-00902540.
XX 10-JUL-2000; 2000US-0217883P.
XX (MONS) MONSANTO TECHNOLOGY LLC.
XX Goldman BS, Hinkle GJ, Slater SC, Wiegand RC;
XX WPI; 2005-028716/03.
XX New substantially purified Myxococcus xanthus nucleic acid molecule
PT encoding a nitrite reductase, useful for determining gene expression,
PT identifying mutations in a gene of interest, and for constructing
PT mutations in a gene of interest.
XX
XX Example 2; SEQ ID NO 16463; 25pp; English.
XX
XX The invention relates to a substantially purified nucleic acid molecule
CC encoding a nitrite reductase of SEQ ID NO. 11926. Further disclosed is a
CC recombinant DNA construct for expression of a nitrite reductase gene in a
CC plant cell, and a plant cell comprising the recombinant DNA construct.
CC The nucleic acid is useful for determining gene expression, identifying
CC mutations in a gene of interest, and for constructing mutations in a gene
CC of interest. Sequences given in records for SEQ IDs 9692-16825 represent
CC a group of 7134 Myxococcus xanthus proteins and peptides. Note: The
CC sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from USPTO
XX Sequence 420 AA;

Query Match 10.1%; Score 275; DB 9; Length 420;
Best Local Similarity 23.5%; Pred. No. 5.7e-16;
Matches 114; Conservative 65; Mismatches 179; Indels 128; Gaps 15;
QY 55 PVLIDNLIQIRHTPHNGL---SDIYLLDDPHEALAAARALIESAEHSLDLOYVWRNDI 111
DB 38 PGVSGGLTRYLPRRHGVQGNACQLLDGVEAYPAMLEAIRGARYVMEYTFVSDA 97
QY 112 SGRLLFNLMYLAERGVRVRLLD-----DNNTRGLDLDLLALDSDHPNTEVR 158
DB 98 VGEFLGQALAEAAERGVRVLYDAVGSWTSRRSFFAGLRAGVD-----IR 144
QY 159 LFNPPVLRKWRALGYLTDFFLRMRMINKSFTADNRATILGGRNIGDEYFKVGEDTVFAD 218
DB 145 AFKPFSLUS--RGLRHLL-----RRDRKILVGDGEVAFVGGVNSIAHWAPAEWGAARD 196
QY 219 LDILATGSVVEGVSHDFDRYWASHSAHNATRIISGNIGKGLQALGYNDTSRHALLYR 278
DB 197 DVLRIGEPAVHELCESATW-----RMMFQGRF-----HRLTRL 232
QY 279 ETVEQSPLYQKIQTGRIDWQSVOTRLISDTPAKG-----LDRDRKPPPIAGRLQDALKOP 333
DB 233 ERLRNP-----PRRGAVGLVLSRRS--IHRAYLHAIRRA 267
QY 334 EKSVLVSPFVPTKSTGDMALAKLVQGDIDVTVLTNSLQATDVAHVSGVVKRKLPLKA 393
DB 268 RRSVLVAAAYFIPDRRWMALEAARRGVEVHLLNA--RSDHPILFEFMAFAVERLLGA 325
QY 394 GIKLYELQPNHVAVPATKDKLTGSSVTSLSHAKTFVQDKRIFTGSENLDRPSARLNTMG 453
DB 326 GVRIFEWQ-----RGV-----LHAKTAVDGVWGTGSGFNLERLSLAFNHEVN 368
QY 454 VVIESPKIAEQMBERTLADTTPEYAYRVTLDKHNRQLQWHDPAKTRKTYPNEPEAKLWKRIAA 513

DB 369 AVFADPRLGQQLDESRGDCGD-CREVTLAEFR-----RPLWQKLE 410
QY 514 KILSLL 519
DB 411 RALSLL 416
RESULT 32
AAW55249
ID AAW55249 standard; protein; 206 AA.
XX AC AAW55249;
XX 26-JUN-1998 (first entry)
XX H. pylori ORF 05cpl191lorf35 protein.
XX Cytoplasmic; vaccine; prevention; treatment; infection; envelope;
KW identification; binding compound; bacteria; life cycle; activator;
KW inhibitor; duodenal ulcer disease; chronic gastritis; diagnosis;
KW bacterium.
XX Helicobacter pylori.
OS
XX WO9737044-A1.
XX 09-OCT-1997.
XX 27-MAR-1997; 97WO-US005223.
XX 29-MAR-1996; 96US-00625811.
PR 02-APR-1996; 96US-00758731.
PR 23-OCT-1996; 96US-00736905.
PR 28-OCT-1996; 96US-00738859.
PR 06-DEC-1996; 96US-00761318.
XX (ASTR) ASTRA AB.
XX Smith D, Alm RA;
XX WPI; 1997-503122/46.
DR N-PSDB; AAV24658.
XX Helicobacter pylori nucleic acid sequences and encoded polypeptide(s) -
PT useful in vaccines to treat or prevent H. pylori infection and for
PT diagnosis of H. pylori infection.
XX Claim 14; Page 492; 1145pp; English.
XX This sequence is a Helicobacter pylori protein of unspecified function.
CC The protein may be used in a vaccine to prevent or treat H. pylori
CC infection or to identify H. pylori polypeptide binding compounds, useful
CC as potential H. pylori life cycle activators or inhibitors. The DNA and
CC probes derived from it may be used for the identification of H. pylori in
CC a sample, and the diagnosis of H. pylori infection. Nucleic acid
CC sequences complementary to the DNA act as antisense sequences, and can be
CC used to prevent the translation of H. pylori mRNA. Antibodies against the
CC protein can be used in immunoassays to evaluate the abundance and
CC distribution of H. pylori-specific antigens. The genomic sequence of H.
CC pylori (ATCC 55679) was determined from overlapping contigs generated by
CC mechanically shearing the bacterial DNA. The sequences were analysed for
CC ORF of at least 180 nucleotides, and the predicted coding regions defined
CC by computer evaluation. To identify likely H. pylori antigens for vaccine
CC development, the amino acid sequences predicted from various ORF were
CC analysed for significant homology to other known or exported membrane
CC proteins. Having identified and determined the sequences of interest,
CC particular regions can be isolated from H. pylori by PCR amplification
CC for recombinant polypeptide production, e.g. in E. coli hosts
XX Sequence 206 AA;
SQ

Query Match 10.1%; Score 274.5; DB 2; Length 206;

Best Local Similarity 31.4%; Pred. No. 2.1e-16;
Matches 64; Conservative 45; Mismatches 84; Indels 11; Gaps 5;
QY 321 PIAGRLQALQKPEKSVYLVSPYFPTKSGTDALAKLVQDGVDTVLTNSLQATDVAHV 380
D 6 PIKIAFAKALNAKESVFIASSYFIPGKKIMKIFKNQISKGIELNLTNSLSSTDIAIVY 65
QY 381 SGYVYKRPDLKAGIKLYELQPNHAPATKOKGLTGSSVTSIHAFTFVQKRIPIGSGN 440
D 66 GAWERYRNKLVRMGANVYEIRNDFPNQIKGR---FSTKSLHGKTIIVFDALTLGSGN 122
QY 441 LDPRSARLNTMGVVIESPKEAQEMERTLADTPYAVRVTLDKHNRLOWHDPATKTY- 499
D 123 IDPRAYINTSVAFLDFNPSFAKRVLSKDHA-QQSHLVLYRH-RVINE--ATEGIL 178
QY 500 ---PNEPEAKLWKRIAAILSLP 520
D 179 IHEKNSPDTSPFLRIKESKVP 202

RESULT 33

ABM69020
ID ABM69020 standard; protein; 490 AA.

AC ABM69020;

XX 20-NOV-2003 (first entry)

XX Photorhabdus luminescens protein sequence #2117.

XX Antibacterial; fungicide; insecticide; polymorphism; genetic analysis;
KW detection; food; gene expression; plant; animal; microorganism; toxin;
KW antibiotic; biopesticide; virulence factor; disease model; plague;
KW whooping cough.

XX Photorhabdus luminescens.

XX WO200294867-A2.

XX 28-NOV-2002.

XX 07-FEB-2002; 2002WO-IB003040.

XX 07-FEB-2001; 2001FR-00001659.

XX (INSP) INST PASTEUR.

PA (CNRS) CNRS CENT NAT RECH SCI.

XX Duchaud E, Taourit S, Glaser P, Frangeul L, Kunst F, Danchin A;
PI Buchrieser C;

XX WPI; 2003-148459/14.

XX Genomic sequence of Photorhabdus luminescens and encoded polypeptides,
PT useful e.g. as therapeutic antimicrobials and agricultural pesticides.

XX Claim 2; SEQ ID NO 2117; 1205pp; French.

XX The invention relates to the isolation of genes and their encoded
CC proteins from Photorhabdus luminescens. The isolated sequences are
CC sources of probes and primers for detecting the genome of P. luminescens
CC and related species; to study polymorphisms; for gene analysis and for
CC detection/amplification of the genes. Antibodies (Ab) raised against the
CC polypeptides encoded by the genes are used for detection/identification
CC of P. luminescens, e.g. in foods. The genes, proteins, Ab and cells that
CC carry a gene-containing vector are used to select compounds that
CC modulate, regulate, induce or inhibit expression of the genes in plants,
CC animals or microorganisms other than P. luminescens and are able to alter
CC response or sensitivity to toxins and antibiotics produced by P.
CC luminescens. Cells transformed to express the genes are useful for
CC recombinant production of the proteins, particularly toxins and
CC antibacterials useful as insecticides, bactericides and fungicides. The
CC genes, proteins, vectors containing the genes and Ab are also useful

CC therapeutically (to treat microbial infection by bacteria or fungi that
CC are sensitive to P. luminescens-encoded toxins or antibiotics) and as
CC biopesticides. Other uses of the genes and the proteins are as virulence
CC factors and for identifying targets of human diseases for which P.
CC luminescens is a model (particularly plague and whooping cough). This
CC sequence represents one of the isolated P. luminescens proteins

XX Sequence 490 AA;

Query Match 10.1%; Score 274.5; DB 6; Length 490;

Best Local Similarity 23.2%; Pred. No. 8e-16;

Matches 118; Conservative 73; Mismatches 185; Indels 133; Gaps 20;

QY 33 SCSSWLPPLERTSRH-FNTSKPVLNDNILOIHTHNGI-----SDVILDDPHAL 86

D 80 STTW---LEELRKSRIHFATENSEVATPLQL--CERROGLAGVKGKIQLLKTYDHS 134

QY 87 ARAALIESAHSLSLOYYIWRNDISGRLLFNLMVLAARVVRVRLDDNNTRGLDILL 146

D 135 KTIIKDIENAKHSIDMVFIWQTGLVEQVTDALMAAAHRCVRCRIWVDSAGS-----WL 189

QY 147 LALDSHPNI-----EVLFPFVLKRWALGYLTFPFLNRMRMKNKSTADNR 194

D 190 YCRSQYPDIMEQAGIEFVEALQVNI FRLFL-----RMDLRQHRKIILIDDY 236

QY 195 ATILGGRNIGD-EYFK-----VGEDTVFADLDILATGSVGVEVSHDFRYWASHSHNATR 249

D 237 ISYAGSMNVDPFRFKQDAGVQ---WIDIMVRMEGPVTTTGMVIYACDMEWETGORIPP 293

QY 250 IIRSGNIGKGLQALGYNDTSRHALLRYRVEQSPLYQKIQTGRIDWQSVQTRLSIDTP 309

D 294 PSPNNIMPFQASGHT-----

QY 310 AKGLDRRRKPIIAGRLQDALKQPEKSVLYSPYFPTKSGTDALAKLVQDGVDTVLTN 369

D 319 --GFPEEL---IQQSLMTAIFSARKQILITTPYFVPSDDLTAICTAALRGVEVIIIVP 372

QY 370 SLQATDVAAVHSGYVKYKPKLLKAGIKLYELQPNHAPATKOKGLTGSSVTSIHAFTIV 429

D 373 --RRNFFLVRWASRAFFAELLNAGVKIYQF-----EDGL-----LHTKSLIV 413

QY 430 DGKRFIFGSFNLDPRSARLNTMGVVIESPKEAQEMERTLADTPPEY---AYRVTLDKHN 486

D 414 DGQLSMVGSVNLDMRSLWLNFETIVVIDDENFGNDL-----TLVQVDYIAHSTLLETH- 466

QY 487 RLOWHDPATRKTYPNEPEAKLWKRIAANKI 515

D 467 --EWE-----QRPLWHRIERI 481

RESULT 34

ABO75565

ID ABO75565 standard; protein; 584 AA.

AC ABO75565;

XX 29-JUL-2004 (first entry)

XX Pseudomonas aeruginosa polypeptide #7740.

XX Bacterial infection; Pseudomonas aeruginosa infection; antibacterial.

XX Pseudomonas aeruginosa.

XX US6551795-B1.

XX 22-APR-2003.

XX 18-FEB-1999; 99US-00252991.

XX 18-FEB-1998; 98US-0074788P.

XX 27-JUL-1998; 98US-0094190P.

XX

PA (GENO-) GENOME THERAPEUTICS CORP.
XX Rubenfield MJ, Nolling J, Deloughery C, Bush D;
XX WPI: 2003-615309/58.
XX N-PSDB; ABD09136.
XX Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide,
XX useful as molecular targets for diagnostics, prophylaxis and treatment of
XX pathological conditions resulting from bacterial infection.
XX Disclosure; SEQ ID NO 24311; 455pp; English.
XX
XX The invention relates to Pseudomonas aeruginosa polypeptides and the
XX polynucleotides encoding them. The sequences are useful in diagnosis and
XX therapy of pathological conditions, as molecular targets for diagnostics,
XX prophylaxis and treatment of pathological conditions resulting from a
XX bacterial infection, for evaluating a compound, such as a polypeptide,
XX for the ability to bind a P. aeruginosa nucleic acid, as components of
XX effective antibacterial targets, as targets for antibacterial drugs,
XX including anti-P. aeruginosa drugs, as templates for recombinant
XX production of P. aeruginosa-derived peptides or polypeptides, as target
XX components for diagnosis and/or treatment of P. aeruginosa-caused
XX infection, and in detection of P. aeruginosa sequences or other sequences
XX of Pseudomonas species using biochip technology. Sequences AB067826-
XX AB084396 represent P. aeruginosa polypeptides of the invention. Note: The
XX sequence data for this patent did not form part of the printed
XX specification but was obtained in electronic format from USPTO at
XX seqdata.uspto.gov/sequence.html
XX
XX Sequence 584 AA;
Query Match 9.9%; Score 268.5; DB 7; Length 584;
Best Local Similarity 24.2%; Pred. No. 3.8e-15;
Matches 107; Conservative 59; Mismatches 167; Indels 109; Gaps 14;
QY 44 RTSEHPTNPKVL---LDNLIQIRHTPHNNGLSDIYLLDDPHEALAAALIESAEHSL 100
Db 192 RRADLHFAVSGPVPWRSDN---RFPALLNDGVRFF---PRMLLA-----IEAAQRSI 237
QY 101 DLOYYIWRNDISGRLLFNLMYLAERGVRVRLLLDDNNTGRLDLDLALDSHPNIEVRLF 160
Db 238 ELEYLVEDGHCALFLVALLDARRRGVAVRCLFDGFCGLGSAWQIRLRAGGELRLY 297
QY 161 NPFVLRKRALG---YLTDFPRLNRRMHKSFADNRATILGGRNIGDEYFK-VGEDTVF 216
Db 298 NPL---RWKLTGGNLY-----RDHRKLLLVDRGLGVGGAGITDEFEPVSDVSAW 345
QY 217 ADLDILATGSGVGVSHDFORYWASHSAHNA-----TRIIRSGNIGKGLQALGYND 267
Db 346 REVNVEMDGPVADWAALFERQWLACLEKAWKPREGMTLTRLPPQGAARGLGRVAYAD 405
QY 268 ETSRHALLRYRETVESQSPLYQKIQTGRIDWQSVQTRLISTPAKGLDRDRRKPPIAGRLQ 327
Db 406 -----ARQHRDILQSLVRALNGSRR-----426
QY 328 DALKQPEKSVLYSPYFVPTKSGTDALAKLVQDGIDVTLTNSLQATDVAAVHSGVVKYR 387
Db 427 -----IWLATPFLPTWKVRRLRKAQGVREVRLLLAG-RLTDHAPVYAGQRY 476
QY 388 KPLLKAGIKLYELQPNHNAVPAKDKIGTSSVTSLSHAKTFIVDGKRFITGSEFNLDPRSAR 447
Db 477 PRLLRAGVRIHEYQPRF-----LHLKVMVMD-DWVSGSCNCFDHNLR 518
QY 448 LNTMGVVIESPKEIAEQMERTL 469
Db 519 FNLDANLEALDPDPTNEAASL 540
RESULT 35
AB068379
ID ABO68379 standard; protein; 579 AA.
XX

AC ABO68379;
XX 29-JUL-2004 (first entry)
XX Pseudomonas aeruginosa polypeptide #554.
XX Bacterial infection; Pseudomonas aeruginosa infection; antibacterial.
XX Pseudomonas aeruginosa.
XX US6551795-B1.
XX 22-APR-2003.
XX 18-FEB-1999; 99US-00252991.
XX 18-FEB-1998; 98US-0074788P.
XX 27-JUL-1998; 98US-0094190P.
XX (GENO-) GENOME THERAPEUTICS CORP.
XX Rubenfield MJ, Nolling J, Deloughery C, Bush D;
XX WPI: 2003-615309/58.
XX N-PSDB; ABD01950.
XX Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide,
XX useful as molecular targets for diagnostics, prophylaxis and treatment of
XX pathological conditions resulting from bacterial infection.
XX Disclosure; SEQ ID NO 17125; 455pp; English.
XX
XX The invention relates to Pseudomonas aeruginosa polypeptides and the
XX polynucleotides encoding them. The sequences are useful in diagnosis and
XX therapy of pathological conditions, as molecular targets for diagnostics,
XX prophylaxis and treatment of pathological conditions resulting from a
XX bacterial infection, for evaluating a compound, such as a polypeptide,
XX for the ability to bind a P. aeruginosa nucleic acid, as components of
XX effective antibacterial targets, as targets for antibacterial drugs,
XX including anti-P. aeruginosa drugs, as templates for recombinant
XX production of P. aeruginosa-derived peptides or polypeptides, as target
XX components for diagnosis and/or treatment of P. aeruginosa-caused
XX infection, and in detection of P. aeruginosa sequences or other sequences
XX of Pseudomonas species using biochip technology. Sequences AB067826-
XX AB084396 represent P. aeruginosa polypeptides of the invention. Note: The
XX sequence data for this patent did not form part of the printed
XX specification but was obtained in electronic format from USPTO at
XX seqdata.uspto.gov/sequence.html
XX
XX Sequence 579 AA;
Query Match 9.9%; Score 268; DB 7; Length 579;
Best Local Similarity 24.5%; Pred. No. 4.1e-15;
Matches 111; Conservative 60; Mismatches 186; Indels 96; Gaps 13;
QY 37 WLPLEERETSESHFTSKPVLDDNLIQIRHTP--HNGLSDIYLLDDPHEALAAALIE 94
Db 186 WRPWEELAAARQVSKYG--LKALVTRTRTTLANN---RVRLJNGEASFEMFKAIS 240
QY 95 SAHSLSLDQYYIWRNDISGRLLFNLMYLAERGVRVRLLLDDNNTGRLDLDLALDSHPN 154
Db 241 AARQVTLVQFFIVRDDALGRLQLLLERAANGVEVFYDAIGSHALPHRYVVERLRQGG 300
QY 155 IEVRLNPNPVLKRWALGYLTDFPRLNRRMHKSFADNRATILGGRNIGDEYFKVGEDT 214
Db 301 VQMHGFS-----TGSGMLNRF-QVNFNHRKVVVVDGECGFGVGHNVGVEY--LGEKP 350
QY 215 VFA---DLDLATGSGVGVSHDF--DRYWASHSAHNAATRIIRSGNIGKGLQALGYND 269
Db 351 PLAPWRDTHMELRGFAVACLQESFADWTWATHSL-----385
QY 270 SRHALLRYRETVESQSPLYQKIQTGRIDWQSVQTRLISTPAKGLDRDRRKPPIAGRLQ 329
XX

vector; a cell comprising a recombinant expression vector of (1);
producing an S. epidermidis polypeptide; an isolated nucleic acid
comprising a nucleotide sequence of at least 8 nucleotides in length; a
vaccine composition for prevention or treatment of an S. epidermidis
infection, comprising a nucleic acid cited above and a carrier; treating
a subject for S. epidermidis infection; a recombinant or substantially
pure preparation of an S. epidermidis polypeptide or its fragment; a
vaccine composition for prevention or treatment of an S. epidermidis
infection; detecting the presence of a Staphylococcus nucleic acid in a
sample; a computer readable medium having recorded in it the nucleotide
sequences with SEQ ID NO: 1-3772 or its fragments; a computer based
system for identifying fragments of the Staphylococcus genome of
commercial importance; a computer based system for identifying fragments
of the Staphylococcus plasmids of commercial importance; identifying
commercially important nucleic acid fragments of the Staphylococcus
genome and/or plasmids; and identifying an expression modulating fragment
of the Staphylococcus genome and/or plasmids. The methods and
compositions of the present invention are useful for the diagnosis,
prevention and/or treatment of an Staphylococcal epidermidis bacterial
infection. This is the amino acid sequence of a S. epidermis protein of
the invention.

XX SQ Sequence 494 AA;

Query Match 9.5%; Score 259.5; DB 8; Length 494;
Best Local Similarity 22.3%; Pred. No. 1.9e-14;
Matches 101; Conservative 70; Mismatches 178; Indels 103; Gaps 11;
QY 76 IYLLDDPHEALAAALIESAEHSLDLOYIWRNDISGRLLFNLMLAERGVRLLLD 135
Db 133 IDLFTDGHKLVKEVLEDIYNADQYIHEYYTFELDGLGKRLDALKKEGLEVKLLYD 192
QY 136 DNNTRGLLDLLALDSDHNIENRVLFPNPFVKRWALGYLTD--PP-----RLNRRMH 185
Db 193 DVGSK-----KVLRSK---FKHFRALGGEVAFPFKVPFLNFRNNRNH 234
QY 186 NKSFTADNRATILGRNIGDYFVKVGEDTVFADLILATGSVVGGEVSHDFRYWASHAH 245
Db 235 RKIIIDQIGYVGFGVGGDYGLGKLGWYRDTHTRVQGCIDALQLRFLDWNQS-- 292
QY 246 NATRIISGNIGKLGALGYNDETSRHALLRYETVEQSPYKIQIOTGRDWSQVTRLI 305
Db 293 -----HRPQFKPDQKYFPKQNG--DKGNAAIQIA 319
QY 306 SDTPAKGLDRDRKPPIAGRLQDALQKQSVVLVSPVFPVTKSGTDALAKLVODGIDVT 365
Db 320 SSGAFDLHQ-----IEYGYTKMIMSAKSIYLOSFPYIPDQSYIINALKWAANSQVEVN 373
QY 366 VLTNLSQATDVAHVHSGYVYKRPKPLKAGIKLYELQPNHAPATKDKGLTGSSTSLHAK 425
Db 374 LMIPC--KPDHPFYVWATFSNAADLLDSGVNIYTYQNGF-----IHSK 414
QY 426 TFIVDGKRIFIGSNLDRSARLNTGEMVIESPKIAEQMERTLADTTPEYAYRVTLDKH 485
Db 415 ILMIDDEISSIGSANMDFRSELPFNFAFYDDEIAKQLRQ-----AFEKDEQS 465
QY 486 NRLQWHPATRKTYPNPEAKLWKRIAALKIS 517
Db 466 KLL-----TKEVDKRPLSIKFKEGLAKLIS 491

RESULT 38
ADA36341
ID ADA36341 standard; protein; 166 AA.
XX
AC ADA36341;
XX
XX 20-NOV-2003 (first entry)
DT
XX Acinetobacter baumannii protein #3502.
DE
XX Acinetobacter baumannii; bacterial disease; antibacterial; vaccine;
KW plant biocontrol agent.

XX Acinetobacter baumannii.
OS
XX US6562958-B1.
PN
XX 13-MAY-2003.
PD
XX 04-JUN-1999; 99US-00328352.
PF
XX 09-JUN-1998; 98US-0088701P.
PR
XX (GENO-) GENOME THERAPEUTICS CORP.
PA
XX Breton G, Bush D;
PI
XX WPI; 2003-576092/54.
DR
XX N-PSDB; ADA32215.
DR
XX New Acinetobacter baumannii proteins and nucleic acids, useful as reagents
PT for diagnosing a bacterial disease, as components of antibacterial
PT vaccines, as targets for antibacterial drugs, or as biocontrol agents for
PT plants.
XX
XX Example; SEQ ID NO 7628; 328pp; English.
PS
XX The invention relates to isolated Acinetobacter baumannii nucleic acids.
CC The A. baumannii nucleic acids and polypeptides are useful as reagents.
CC for diagnosing a bacterial disease, as components of antibacterial
CC vaccines, as targets for antibacterial drugs, to detect the presence of
CC A. baumannii and other Acinetobacter species in a sample, in screening
CC compounds for the ability to interfere with the A. baumannii life cycle
CC or to inhibit A. baumannii infection, and as biocontrol agents for
CC plants. The present sequence represents the amino acid sequence of an A.
CC baumannii protein.
XX
SQ Sequence 166 AA;
Query Match 9.3%; Score 252; DB 6; Length 166;
Best Local Similarity 37.3%; Pred. No. 1.7e-14;
Matches 60; Conservative 26; Mismatches 51; Indels 24; Gaps 4;
QY 382 GYVKRXPLLKAGIKLYELQ-----NHAVPATKDKGITGSSVTSLSHAKTF 427
Db 13 GVRQYRVEMLKSGIELVEFKPVLERRRRRTWVEIVTGSVIPA-----KGKXKSLHAKFF 66
QY 428 IVDGKRIFIGSNLDRSARLNTGEMVIESPKIAEQMERTLADTTPEYAYRVTLDKHNR 487
Db 67 DVDGK-VFIGSFNDFRSTYLNTEVGLVIESQLQTSVMLDQHLPOVAYQLKNSGQ 125
QY 488 LOWHDPATRK--TYNPEAKLWKRIAALKISLLPIEGLL 525
Db 126 ITWLDYQANGQVIEYDKDPTSRFQRTWIKAVSYLPIEWMM 166
RESULT 39
AAB76548
ID AAB76548 standard; protein; 422 AA.
XX
AC AAB76548;
XX
XX 11-APR-2001 (first entry)
DT
XX Corynebacterium glutamicum MCT protein SEQ ID NO:78.
DE
XX Corynebacterium glutamicum; brevivacterium lactofermentum; MCT;
KW membrane construction and membrane transport protein; petroleum spill;
KW hydrocarbon degradation; gram positive aerobic bacterium; marker;
KW identification; microorganism; fine chemical production; transformation;
KW genome mapping; genetic engineering.
XX
OS Corynebacterium glutamicum.
XX
XX WO200100805-A2.

are useful for identifying the mutation point of a gene derived from a mutant of coryneform bacterium, measuring expression amount and analysing the expression profile or expression pattern of a gene derived from Coryneform bacterium, and identifying a homologue of a gene derived from coryneform bacterium. Coryneform bacteria are useful for producing amino acids, nucleic acids, vitamins, saccharides and organic acids, particularly L-lysine. The present sequence is a protein described in the exemplification of the invention. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from the European Patent Office

XX Sequence 500 AA;

Query Match		9.3%;	Score 252;	DB 4;	Length 500;
Best Local Similarity		23.4%;	Pred. No. 9.7e-14;		
Matches 108;		Conservative 58;	Mismatches 179;	Indels 116;	Gaps 17;
Qy	67	TPHNGLSDIYLLDDPHEALAAARALIESAFHSLDLOYYIWRNDISGRLLFNLMYLAER	126		
Db	116	TGGNNG-----FYSYRESLKRMTAAIDEAEYIYVEIYIMAWDSYTPFFAALERAHNR	170		
Qy	127	GVRVRLLD-----DNNTRGLDDLLALDSDHPNIEVRLFNPFVLRKWRAL	171		
Db	171	GKVRLLFDHVGSKWYPCYHRLKKELNRMGFAYLM-----LPLQWR--	213		
Qy	172	GYLTDFPRLNRRMHKSFADNRATILGGRN-IGDEY-----PKVGEDTVFADLDILATG	225		
Db	214	---RRFRPDLNRHKMLIIDGHTAFNGSQNLTAPSYLQKNYKIGRE--WKDLMWELTG	268		
Qy	226	SVGVEVSHDFRYWASHSAHNATRIIRSGNIGKGLQALGYNDETSRHALLRYRETVEQSP	285		
Db	269	PIVSSMEMIFAGDWYVESN-----EALDIRDHAEAHGYIGNTQKOSATN	312		
Qy	286	LYQKIQTGRIDWQSVQTELLISDTFPAKGLDRDRKPPIAGRLQDALKQPEKSVILVSPYFV	345		
Db	313	LVQLIPSG-----PGYTEPNLR-----MFNSIVHHAKERLILCSPYFI	351		
Qy	346	PTKSGTDALAKLVODGIDVTVLNLSLOATDVAAVHSGYVKYKPLLKAGIKLYELOPNHA	405		
Db	352	PDESLLLEAVTSACYRGVTVELFVS--EQADQFAIDHAQSSYYQALLEAGVKIYQF-PKPD	408		
Qy	406	VPATKDKGLTGSSVTSLHAKTFIVD-----GKRIF--IGSFNLDPRSARLNTMGVVIES	458		
Db	409	V-----LHTKYMIAADPDDTTGNEALGVLGSSNLDIRSFGLNVEISLMIK	453		
Qy	459	PKIAEQMERTLADTTPEYAYRVTLDKHNRLOWHDPATKTY	499		
Db	454	GNIHEL-NALTDRTYTVSFKLTDKWNQRSW-----RRRY	488		

Search completed: May 2, 2006, 06:09:19
Job time : 193 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: May 2, 2006, 06:13:30 ; Search time 46 Seconds
(without alignments)
943.581 Million cell updates/sec

Title: US-10-665-990A-14
Perfect score: 2720
Sequence: 1 MHTDPKIQAMPSETISPMKT.....KLMKRIAATLPIEGLL 525

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

- 1: /cgn2_6/ptodata/1/iaa/5_COMB.pep.*
- 2: /cgn2_6/ptodata/1/iaa/6_COMB.pep.*
- 3: /cgn2_6/ptodata/1/iaa/H_COMB.pep.*
- 4: /cgn2_6/ptodata/1/iaa/PCITUS_COMB.pep.*
- 5: /cgn2_6/ptodata/1/iaa/RE_COMB.pep.*
- 6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	831.5	30.6	543	2	US-09-540-236-2405
2	790	29.0	564	2	US-09-252-991A-23909
3	493.5	18.1	271	2	US-09-328-352-5559
4	478.5	17.6	275	2	US-09-328-352-5591
5	409	15.0	321	2	US-09-328-352-7685
6	291	10.7	495	2	US-09-134-001C-4273
7	286.5	10.5	493	2	US-09-828-523A-22
8	286.5	10.5	502	2	US-09-828-523A-90
9	275	10.1	420	2	US-09-902-540-16463
10	268.5	9.9	584	2	US-09-252-991A-24311
11	268	9.9	579	2	US-09-252-991A-17125
12	259.5	9.5	494	2	US-09-134-001C-4176
13	252	9.3	166	2	US-09-328-352-7628
14	252	9.3	422	2	US-09-602-787A-78
15	252	9.3	500	2	US-09-848-726-2
16	246	9.0	508	2	US-09-543-681A-4717
17	241	8.9	409	2	US-09-134-000C-3664
18	232	8.5	442	2	US-09-710-279-1000
19	231.5	8.5	386	2	US-09-107-532A-4088
20	228	8.4	800	2	US-09-252-991A-24635
21	220	8.1	501	2	US-09-489-039A-10241
22	219	8.1	491	2	US-09-583-110-4158
23	216	7.9	415	2	US-09-134-000C-3595
24	215.5	7.9	419	2	US-09-489-039A-12789
25	215	7.9	518	2	US-09-107-433-3177
26	208	7.6	378	2	US-09-902-540-11860
27	149	5.5	213	2	US-09-902-540-13301

28	141.5	5.2	478	2	US-09-438-185A-437	Sequence 437, Appl
29	141.5	5.2	933	2	US-09-141-206-2	Sequence 2, Appli
30	135.5	5.0	933	2	US-09-949-016-9934	Sequence 9934, Ap
31	134	4.9	922	2	US-09-141-206-6	Sequence 6, Appli
32	133.5	4.9	458	2	US-09-489-039A-11632	Sequence 11632, A
33	133.5	4.9	932	2	US-08-968-752B-6	Sequence 6, Appli
34	133.5	4.9	932	2	US-09-536-224-6	Sequence 6, Appli
35	133.5	4.9	933	2	US-09-107-149-17	Sequence 17, Appl
36	130.5	4.8	272	2	US-09-902-540-16500	Sequence 16500, A
37	130	4.8	459	2	US-09-543-681A-5540	Sequence 5540, Ap
38	123	4.5	155	2	US-09-902-540-14817	Sequence 14817, A
39	122	4.5	183	2	US-09-107-532A-5960	Sequence 5960, Ap
40	121.5	4.5	933	2	US-09-107-149-2	Sequence 2, Appli
41	119	4.4	117	2	US-09-328-352-7334	Sequence 7334, Ap
42	111.5	4.1	961	2	US-09-328-352-7358	Sequence 7358, Ap
43	110	4.0	1036	2	US-08-968-752B-4	Sequence 4, Appli
44	110	4.0	1036	2	US-09-536-224-4	Sequence 4, Appli
45	109	4.0	510	2	US-09-540-236-3660	Sequence 3660, Ap

ALIGNMENTS

RESULT 1

US-09-540-236-2405
; Sequence 2405, Application US/09540236
; Patent No. 6673910
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO MORAXELLA CAT
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2005-001
; CURRENT APPLICATION NUMBER: US/09/540, 236
; CURRENT FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 3840
; SEQ ID NO 2405
; LENGTH: 543
; TYPE: PRT
; ORGANISM: M.catarrhalis
US-09-540-236-2405

Query Match 30.6%; Score 831.5; DB 2; Length 543;
Best Local Similarity 35.9%; Pred. No. 2.1e-74;
Matches 194; Conservative 102; Mismatches 198; Indels 47; Gaps 12;

Qy	22	SLISLCLLCSSWLP-----PLEERTSRHFNTPSKVLLDNIL-----QIRHT	67
Db	13	SIHTVMLVLLVGCKN-LPDTPLPKSLALTQVQRHQVMYKGGSSDGLISAINAQSHHH	71
Qy	68	PHNGLSDIYLLDDPHEALAAARALTESAEHSDLOYYIWRNDISGRLLFNLMVLAARG	127
Db	72	PNOSG---YYPITTGANAFAARSTLTDVANQSDIQYIWHNDEAGOLMKDLWEADRG	128
Qy	128	VRVRLDD-NNTRGLDDLLALDSHPNIEVRLNPFVLRKWRALGYLTDFPRNRRMHN	186
Db	129	VIVRLDDFNSPELDQLLRISKKNIAVRLINPMPYRGFRSLNMLHPITRNRRMHN	188
Qy	187	KSTADNRATILGRRNIGDEYFKVGEDTFADLDILATGSGVGEVSHDFRYWASHAHN	246
Db	189	KSMTFDNKISVIGRRNIGNEYNNAFVNHFADLDVMLGVHVGKITQSFYIYASPLSFD	248
Qy	247	ATRIIRSGNIG--KGQAALYND-----ETSRHALLRYRETVEOSPLQYQKIQOTGRIDW	297
Db	249	IELVLRHNDKDDISGKPVVDFDELEKVKOSSDADRELRTYRQAMQNSTIQDILLAQQVPF	308
Qy	298	QSVQTRLI SDTPAKGLDRDRRKPPPIAGRLQDALKQPEKSVYLVSVYFVPFKSGTDLAKL	357
Db	309	FWTEIDLADNVIKLNGHSDPNEFLVSQLTQLGQPSKKLSIISYFVPTKEGIDTLITL	368
Qy	358	VQDGIDVTILNLSQATDVAAVHSGYVKYKPKLLKAGIKLYELOPN-HAVPATKDK--GL	414
Db	369	AKMGVEVRLTNSFDTADTVGIHVAGYAHWRKQLLAAGVHLFEIKSSAQSIQDNENRFWT	428


```
Qy 33 SCSSWLPPLERTSRH-FNTSKPVLNLIQIRHTPHNGL-----SDIYLLDDPHEAL 86
| : ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 99 SWAW- --LEBLKCKHIFAKHYSVAEPIQL--TAKRQINGVKGKIKIQLLTCCDSL 153
| : ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 87 AARALTESAHSLLDQYIWRNDISGRLLFNL--MYLAERGVVRVRLDLDNNTGLD 143
| : ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 154 NAITRDINNARDNIEMVFIWQ---SGGLVEEVEALIQAAKRGVKKIMVDSAGSRFF 210
| : ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 144 DL-----LLALDHPNIEVRLFPVLRKWRALGYLTFPRLNRRMHNKSFTADNRAT 196
| : ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 211 RTNGPARMAAGIEFVESLQVNLFR-FELR-----RMDLROHRKIVLIDNVIS 257
| : ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 197 ILGRNIGD-BYFK-----VGEDTFADLDIATGSV---VGEVSHDFPRYWASHAHNAT 248
| : ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 258 YTGSMNMVDPFYKQDAGVGQ---WIDIMRMGPGVSTTLG-IIVAFD--WEMETGERHL 311
| : ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 249 RIIRSGNIGKGLQALGYNDETSRHALLRYRETVEQSPLYQKIOTGRIDWQSVOTRLSDT 308
| : ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 312 PPPDNNIMPEQANGHT-----TQVIASGPGFPE-----ELIQOSLIT-- 350
| : ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 309 PAKGLDRDRRKPPIAGRLQDALKOPEKSVYLVSPYFVPTKSGTDLAKLVDQDGVTLT 368
| : ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 351 -----AIYSARKELVMTTFYFVPSDDLAAHAISTAAMRGVDVSIIV 390
| : ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 369 NSLQATDVAHVHSGVYKVKPLLKAGIKLYELQPNHAVPATKGLTGSSVTSLHAKTFI 428
| : ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 391 P--RSNDSFLVRWASRAFFTILEAGVKVQF-----EDGL-----LHTKSVM 431
| : ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 429 VDGKRIFTGSFNLPDSARLANTMGVIESPKIAEQMERTLADTTPEYAY-----RVTL 483
| : ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 432 VDGQSMVGSVNLDMRSLWLNFEITVVIDBCFASDL-----SIVQDYIARSTOLTLD 485
| : ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 484 K 484
| : ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 486 E 486

RESULT 17
US-09-134-000C-3664
; Sequence 3664, Application US/09134000C
; Patent No. 6617156
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; FILE REFERENCE: ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 032796-032
; CURRENT APPLICATION NUMBER: US/09/134,000C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/055,778
; PRIOR FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 6812
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3664
; LENGTH: 409
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
US-09-134-000C-3664
```

```
Query Match 8.9%; Score 241; DB 2; Length 409;
Best Local Similarity 23.8%; Pred. No. 3.1e-15;
Matches 107; Conservative 60; Mismatches 188; Indels 94; Gaps 15;

Qy 40 PLEERTESRHFNTS---KPVLLDNIQIRHTPHNGLSDIYLLDDPHEALAAALIESA 96
| : ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 19 PTEEQQQTDFESSLNRMPLTKN-----DVEIFTDGTKEFNALMADIKKA 64
| : ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 97 EHSLLDQYIWRNDISGRLLFNLMYLAERGVVRVRLDLDNNTGLDILLALDHPNIE 156
| : ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 65 QHSIHIEYAFVTHIGTKIINLLEKAAEGVEVELLYDAFGSKG-----TKVHILNE 117
| : ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 157 VRLFPVLRKWRALGYLTFPRLNRRMHNKSFTADNRATILGGRNIGDEYFKVGEDTVF 216
| : ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
```

```
Db 118 LKXNGGFVQTFITSQKALLKF-RLNYHDHRKIVVIDGKVGIVGGFNVDQY--AGTTKKF 174
| : ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 217 ADLDILATGVSVEVSHDFDRYWASHAHNATRIIRSGNIGKGLQ---ALGYNDETSRHA 273
| : ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 175 G-----YWRD-----THLRIQCPAASLLQMRFLMDWNVSSPEKN 208
| : ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 274 LLRYRETVEQSPLYQKIOTGRIDWQSVOTRLSDTPAKGLDRDRRKPPIAGRLQDALQKP 333
| : ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 209 RVAY-----QLDYEFKLE-ALVPEANTSIQMI-----ASGPNSDREQIKLA--FIKLITSA 256
| : ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 334 EKSIVLVSPYFVPTKSGTDLAKLVDQDGVTLVLTNSLQATDVAHVHSGVYKVKPLLKA 393
| : ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 257 KKRWIQTPVLPVDPDSVLAAALKVAASGVVDKIMP--DKPDHFFIYRATQYVYGRLLMKE 314
| : ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 394 GIKLYELQPNHAVPATKGLTGSSVTSLHAKTFIVDGKRIFTGSFNLPDSARLANTMG 453
| : ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 315 NIEILVYNGGF-----LHAKTWMIDDEVCTGVSANQDIRSYKLNPEAN 357
| : ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 454 VVIESPKIAEQME-----RTLADTTPE 475
| : ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 358 AVLYDKKIIDQLEAIFLEDRKKCTTWTPE 386
| : ||| : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 18
US-09-710-279-1000
; Sequence 1000, Application US/09710279
; Patent No. 6703492
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PU3480US
; CURRENT APPLICATION NUMBER: US/09/710,279
; CURRENT FILING DATE: 2000-11-09
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1000
; LENGTH: 442
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: amino acid sequence
; FEATURE:
; NAME/KEY: MOD RES
; LOCATION: (442)
; OTHER INFORMATION: variable amino acid
US-09-710-279-1000
```

```
Query Match 8.5%; Score 232; DB 2; Length 442;
Best Local Similarity 22.8%; Pred. No. 2.8e-14;
Matches 89; Conservative 57; Mismatches 157; Indels 88; Gaps 9;

Qy 76 IYLLDDPHEALAAARALIESAHSLLDQYIWRNDISGRLLFNLMYLAERGVVRVRLDLD 135
| : ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 129 IDLFTDGHKLYEKVLEDIYNAQDYIHLEYITFELDGLGKRLDQALETKLKGLEVLKLYD 188
| : ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 136 DNNTRGGLDILLALDHPNIEVRLFPVLRKWRALGYLTD--FP-----RLNRRMH 185
| : ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 189 DVGSK-----KVRLSK---FKHFRALGGEVEAFPPSKVPLINFRMNNRNH 230
| : ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 186 NKSTANDRATILGGRNIGDEYFKVGEDTVFADLDILATGVSVEVSHDFPRYWASHAH 245
| : ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 231 RKIIIDQIGYGVGVNVDYLGILGLGYWRDTHTRVQGEICDALQRLFLDWNSSQ-- 288
| : ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 246 NATRIIRSGNIGKGLQALGYNDETSRHALLRYRETVEQSPLYQKIOTGRIDWQSVOTRL 305
| : ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 289 -----HROPKFDQKFFKKG--DKGNAAIOIA 315
| : ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 306 SDTPAKGLDRDRRKPPIAGRLQDALKOPEKSVYLVSPYFVPTKSGTDLAKLVDQIDVT 365
| : ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 316 SSGPAFDLHQ-----IEYGYTKMIMSAAKSIYLSQSPYFIPDQSYINALKMAANSQVGN 369
| : ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
```

Qy 366 VLTNLSQATDVAAGHYKVRKPLKAGIKYELQPNHVAHPATKDKGLTGSSVTSLHAK 425
Db 370 LMIPK--KPDHPFVYVATFNAADLDSGVNIYYQNGF-----IHSK 410
Qy 426 TFIVDGKRIFGTSFNLDPRSARLNTMGVWI 456
Db 411 ILMIDDEISSIGSANMDFRSFLNFEVNAFI 441

RESULT 19

US-09-107-532A-4088
; Sequence 4088, Application US/09107532A
; Patent No. 6583275
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 7310
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354

; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: PC
; OPERATING SYSTEM: <Unknown>
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,532A
; FILING DATE: 30-Jun-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/085,598
; FILING DATE: 14 May 1998
; APPLICATION NUMBER: 60/051571
; FILING DATE: July 2, 1997

; ATTORNEY/AGENT INFORMATION:
; NAME: Ariniello, Pamela Deneke
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GTC-012
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781)893-5007
; TELEFAX: (781)893-8277
; INFORMATION FOR SEQ ID NO: 4088:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 386 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHEICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: Enterococcus faecium

; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (B) LOCATION 1...386
; SEQUENCE DESCRIPTION: SEQ ID NO: 4088:
US-09-107-532A-4088

Query Match 8.5%; Score 231.5; DB 2; Length 386;
Best Local Similarity 23.3%; Pred. No. 2.5e-14;
Matches 97; Conservative 66; Mismatches 174; Indels 79; Gaps 14;
Qy 58 LDNILOIRHTPHNGLSDIYLLDDPHEALAAARALIESAHSLLDQYIWRNDISGRLLF 117
Db 2 LATYKICPAQTR--MNCVTLTLDGQEKLLSLQDLKAKKHSIHIEYFAEVTNIGQQVL 60

Qy 118 NMYLAAERGVRVRLLDNDNTRGLDILLALDHPNIEVRLNPFVLRKWRALGYLTD 177
Db 61 HVLEKAAEGVEVRILYD-----AFGSH-GTKAKDFNRLLKNGHVVHTFVTSQ 107

Qy 178 P-----RLNRRMHNKSFADNRATILGGRNIGDEYFKVGEDTVFADLDILATGSSVGVBS 232
Db 108 RALLRFLNVDHDKIIVVIGDKISYTGGFNIANY-----VNTTKFGYWR 153
Qy 233 HFDPRYWAHSAHNATRIIRSGNIG-KGLQALGYNDETSRHALLRYRETVEQSPLYQKIQ 291
Db 154 DTHIRIFCAASSLLQLRFLTDMNVSVPEEKGVYH-----LNY-----FFKKAD 197
Qy 292 TGRIDMQSVQTRLSIDTPAKGLDRDRPKPIAGRLQDALQKPEKSVLVSPYFVPTKSGT 351
Db 198 RDESKLADTSLQVSSGPN--EREQIKLSFKLITS-----KRVMIQTPLYVPDESVI 251
Qy 352 DALAKLVQDGDIVTLTNSLQATDVAAGHYKVRKPLKAGIKL--YELQPNHVAHPAT 409
Db 252 AALKIATASGVGVKIMIPN--KPDHPFIYRATQYARQLIKENVOILVYE----- 299
Qy 410 KDGLTGSSVTSLHAKTIFVIGDKRIFGTSFNLDPRSARLNTMGVWIESPKIABQM 465
Db 300 --NGF-----LHAKTLIMDDEICMVGSANQDIRSYRLNPFETSASIVYDPEFLEL 346

RESULT 20

US-09-252-991A-24635
; Sequence 24635, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 24635
; LENGTH: 800
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-24635
Query Match 8.4%; Score 228; DB 2; Length 800;
Best Local Similarity 23.2%; Pred. No. 1.9e-13;
Matches 103; Conservative 61; Mismatches 186; Indels 94; Gaps 14;
Qy 68 PHNGLSDIYLLDDPHEALAAARALIESAHSLLDQYIWRNDISGRLLFNLMLYAAERG 127
Db 72 PWRDG-NRVELLNGEYFPRLFCIAEARREILLETFFIFEDEVGRQLQEAALSAARG 130
Qy 128 VYRLLLDNDNTRGLDILLALDHPNIEVRLNPFVLRKWRALGYLTDPRNRRMHNK 187
Db 131 VEQVTVGVGTASLSPDYLAARLAASGVRVHLFDPP-----RPRLMGRTN---LFRRLHRK 183

Qy 188 SFTADNRATILGGRNIGDEYFKVGEDTVFADLDILATGSSVGVSHDFDRYWAHSAHNA 247
Db 184 LVVIDRRQAFVGGINGEDH----- 203
Qy 248 TRIIRSGNIGK---GLQALGYNDETSRHALLRYRETVEQSPLYQKIQOTGRIDMQSVQTRL 304
Db 204 --LVRRGNMAQDYAVRVEGPPVVRDIRQACIALLEPDADYPLRPSGAG---QPARVRL 257
Qy 305 ISDTPAKGLDRDRPKPIAGRLQDALQKPEKSVLVSPYFVPTKSGTDLAKLVQDQIDV 364
Db 258 VI-----RNDQSSDDIEREYLOAIRARRRLLIANAYFFPGYRLLRELDAARRGVRV 311
Qy 365 TVLTNSLQATDVAAGHYGVV-KYRKPLKAGIKLYEL--QPNHVAHPATKDKGLTGSSVTS 421
Db 312 DLVLQGMPPMPLVRLCSRLLYDY---LLREGVRIHEYVQRP----- 349
Qy 422 LHAKTIFVIGDKRIFGTSFNLDPRSARLNTMGVWIESPKIAEQMERTLADTTPEYARVT 481

[illegible]

Query Match	7.9%;	Score 215.5;	DB 2;	Length 419;
Best Local Similarity	21.4%;	Pred. No. 1.2e-12;		
Matches 100;	Conservative 59;	Mismatches 204;	Indels 105;	Gaps 11;
QY	76	IYLLDDPHEALARAALIEGAEHSLDQYIYWRNDISGRLLFNFLMWLAAERGVRVLLLD	135	
Db	16	ITLLENGDQYYPALFAAIGASRRVILIESPTWFEDVGRRLHAVLLEARRGIGQVEVLLD	75	
QY	136	DNNTRGLDLLLLALDSDHNTIEVLFNFVLRKWRALGYLTDFPRLNRRMHMKNSTADNRA	195	
Db	76	GYGSPDLSDFVGLGTAAGVIIFYRDP----RPKLMGMRTN---LFRMRHKIYVVIDDTT	128	
QY	196	TILGGRNIGDEYFKVGEDTVPADLLDILATGVSUVGEVSHDPRYWASHAHNATRIIRSGN	255	
Db	129	AFVGGINYSAEHMSDYGPEAKQDYAVKVEGVPVLLDIL-QFE-----	168	
QY	256	IGKGLQALGYNDETSRHALLRYRETVEQSPDYQKIQTGRIDMQSVQTRFLSDTPAKG---	312	
Db	169	----LENLP-NSSTARWRWRRRHQPE-----INOTGEQAQAL	201	

Qy	313	---	LDRRRKPPIAGRIQDALKQPEKSVYLVSPYFPTKSGTDALAKLVQDGIDVTVLTN	369
Db	202	FVWRDNQDHRRDIERHYLKMLTSAARVEIIANAYFFPGYRLLHAMRNAARRGVRVKLIWQ	261	
Qy	370	SLOATDVAAVHSGVYKRPPLLKAGIKLYELOPHNAVPA TKDKGLTGSSVTSLSHAKTFIV	429	
Db	262	G--EPDIPITVFGAHLIYHVYLKVGGVQIYEYRR-	302	
Qy	430	DGKRIFIGSNLPORSARLNTMGWVIESPKIAQMERTLADTTPEAYRYRTLDKHNRLQ	489	
Db	303	DDHWATVGSNLDPLSLSLNLLEANLIITHDRVFNQTLRDNLNGLTARDQCQRIDKTMLPKRN	362	
Qy	490	W-----HDPA-----TRKTYNEPPEAKLWKRI	511	
Db	363	WRLGVSMAFHLRHFPANVGMPLPAHTPRLARVSPPVPORIEOTDRV	410	

RESULT 25

US-09-107-433-3177
; Sequence 3177, Application US/09107433
; Patent No. 6800744
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID
; SEQUENCES RELATING TO STREPTOCOCCUS PNEUMONIAE
; THERAPEUTICS
; NUMBER OF SEQUENCES: 5206
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Walham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354

..

MEDIUM TYPE: CD/ROM ISO9660
COMPUTER: <unknown>
OPERATING SYSTEM: <unknown>
SOFTWARE: <unknown>
CURRENT APPLICATION DATA:
 APPLICATIION NUMBER: US/09/107,433
 FILING DATE: 30-Jun-1998
PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 60/ 085131
 FILING DATE: May 12, 1998
 APPLICATION NUMBER: 60/051553
 FILING DATE: July 2, 1997
ATTORNEY/AGENT INFORMATION:
 NAME: Ariniello, Pamela Deneke
 REGISTRATION NUMBER: 40,489
REFERENCE/DOCKET NUMBER: GTC-011
TELECOMMUNICATION INFORMATION:
 TELEPHONE: (781)893-5007
 TELEFAX: (781)893-8277
INFORMATION FOR SEQ ID NO: 3177:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 518 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: YES
ORIGINAL SOURCE:
 ORGANISM: Streptococcus pneumoniae
FEATURE:
 NAME/KEY: misc feature
 LOCATION: (B) LOCATION 1...518
SEQUENCE DESCRIPTION: SEQ ID NO: 3177:
US-09-107-433-3177

Query Match	7.9%;	Score 215;	DB 2;	Length 518;
Best Local Similarity	21.2%;	Pred. No. 1.9e-12;		
Matches 100;	Conservative	61;	Mismatches 140;	Indels 170;
			Gaps	20;

! TYPE: PRT
! ORGANISM: Human
US-09-949-016-9934

Query Match 5.0%; Score 135.5; DB 2; Length 933;
Best Local Similarity 18.9%; Pred. No. 0.0005;
Matches 123; Conservative 92; Mismatches 179; Indels 257; Gaps 34;
QY 30 LLCSSS-----WLPPLERTES-----RHFNSTSKP-----VL 57
Db 289 LILKSSYRQARWAQETELAQGPRDFLOLHRHDSYAPPRPGTLARFWVNGAGYFAAV 348
QY 58 LDNILOIRHTPHNGLSDIYLLD---DPHEALAAALIESAEHSLDLQYIWRNDISGR 114
Db 349 ADAILRAQ-----BEIFITDWLSP-EVYLKRP-----HSD--WRLDI--- 385
QY 115 LFNLMYLAERGVVRVRLDDNNTRGLDGLLALDS-----HPNIEVRLNPFV 164
Db 386 ---MLKRAEGGVRSVILLFKE---VELALGINSVSKREALMLLHPNIV-MRHPDQ 435
QY 165 LRKWRALGYLTFPRLNRRMHNKSTADNRATILGG-----RNIGD----- 205
Db 436 VTLW-----AHHEKLLVVDQVVAFLGGLDLAYGRWDDLHYRLTDLGDSSES 481
QY 206 -----EYFKVGEDTFADLDILATGVSVEVSHDF---DRYWAS 241
Db 482 AASQPTPRPDSPATPDLSHNOFFWLGKD---YSNL-----ITKDWQLDRPFED 528
QY 242 HSAHNATRIIRSGNIGKGLQALGYNDTSRHALLRYRETVE-----QSPLYOKIQTGRID 296
Db 529 FIDRETTPRMPWRDGVVHGLPAR-LARHFIQRWNETKTKAKVKTPIYPYLLPKS-- 585
QY 297 WQSVOTRLISDTPA-----KGLDRDRKPPPIAGRLQDA-----LKQPKSYL 339
Db 586 -TSTANQLPFTLPGGQCTTVQVLRSDVRWS-----AGTLENSILNAYLHTIRESQHFLYI 639
QY 340 VSPYFVPTKSGTDALAKLVQDGID-----VTVLT-----NSL 371
Db 640 ENQFISCSGRTVLNKVGEIVDRILKAHKQGCWYRVVLLPLLPFGEGDISTGGNSI 699
QY 372 QATDVAHVHSGY-----VKYRKPLLKAGIKYELQPNHAVPATKDK 412
Db 700 Q-----AILHFTYRTLRCRGEYSILHRLKAAMGTAWRDYISICGLRTHGELGHPV----- 749
QY 413 GLTGSSVTSLHAKTIVDGKRFIFSGFNLDPRS--ARLNTMGVWIE---SPKI---AE 463
Db 750 ----SELIIYHKSVLIAADRVTIIIGSANINDRSLGKRDSSELAVLIEDTETEPSLMNGAE 805
QY 464 QMERTLADTTPYAYRVTLDKHNR--LOWHDPATKTYPNPEAKLWKRIA 512
Db 806 YQAGRFALSLRKHCFCGVILGANTRPDLDRDPICDDFF-----QLWQDMA 850

RESULT 31
US-09-141-206-6
; Sequence 6, Application US/09141206
; Patent No. 6187559
; GENERAL INFORMATION:
; APPLICANT: Steed, Paul M.
; APPLICANT: Lasala, Daniel J.
; TITLE OF INVENTION: Amino Acid Sequence of Human PLD2A
; Patent No. 6187559
; FILE REFERENCE: 4-30148/PI/CGC1954/R
; CURRENT APPLICATION NUMBER: US/09/141,206
; CURRENT FILING DATE: 1998-08-27
; EARLIER APPLICATION NUMBER: 60/057,802
; EARLIER FILING DATE: 1997-08-28
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 922
; TYPE: PRT
; ORGANISM: Human

US-09-141-206-6

Query Match 4.9%; Score 134; DB 2; Length 922;
Best Local Similarity 18.4%; Pred. No. 0.0007;
Matches 121; Conservative 84; Mismatches 173; Indels 278; Gaps 32;
QY 30 LLCSSS-----WLPPLERTES-----RHFNSTSKP-----VL 57
Db 289 LILKSSYRQARWAQETELAQGPRDFLOLHRHDSYAPPRPGTLARFWVNGAGYFAAV 348
QY 58 LDNILOIRHTPHNGLSDIYLLD---DPHEALAAALIESAEHSLDLQYIWRNDISGR 114
Db 349 ADAILRAQ-----BEIFITDWLSP-EVYLKRP-----HSD--WRLDI--- 385
QY 115 LFNLMYLAERGVVRVRLDDNNTRGLDGLLALD-----SHNIEVRLNFP 162
Db 386 ---MFKRAEGGVRSVILL-----FKELEALGINSVSKREALMLLHPNIV-MRHP 433
QY 163 FVLKRWALGYLTFPRLNRRMHNKSTADNRATILGG-----RNIGD----- 205
Db 434 DQVTLW-----AHHEKLLVVDQVVAFLGGLDLAYGRWDDLHYRLTDLGDS 479
QY 206 -----EYFKVGED-----TV 215
Db 480 ESAASQPTPRPDSPATPDLSHNOFFWLGKDYNSLITKDWQLDRPFEDFIDRETTPRMP 539
QY 216 FADLDILATGVSVEVSHDFRYWASHAHNATRIIRSGNIGKGLQALGYNDTSRHALL 275
Db 540 WRDGVVHGLPARDLARHFIQRW-----NFTKTKA-----KYKTPTYPYLLP 583
QY 276 RYRETVESQPLYQIKIQTGRIDWQSVQTRLSIDTPAKGLDRDRKPPPIAGRLQDA----- 329
Db 584 KSTSTANQLPF--TLPGGQC--TTVQV-----LRSDVRWS-----AGTLENSILNAYL 627
QY 330 --LKQPKSYLVGYPYFVPTKSGTDALAKLVQDGID-----VTVLT----- 368
Db 628 HTIRESQHFLYENQFFISCSGRTVLNKVGEIVDRILKAHKQGCWYRVVLLPLLP 687
QY 369 -----NSLOATDVAHVHSGY-----VKYRKPLLKAGIKYEL 400
Db 688 EGDISTGGNSIQ-----AILHFTYRTLRCRGEYSILHRLKAAMGTAWRDYISICGLRTHGE 743
QY 401 QPNHAVPATKDKGLTGSSVTSLHAKTIVDGKRFIFSGFNLDPRS--ARLNTMGVWIES 458
Db 744 LGHPV-----SELIIYHKSVLIAADRVTIIIGSANINDRSLGKRDSSELAVLIED 793
QY 459 PKIAEQMERTLADTTPYAYRVTLDKHNR--LOWHDPATKTYPNPEAKLWKRIA 512
Db 794 ---TETEPSLMNGAEYQAGSVILGANTRPDLDRDPICDDFF-----QLWQDMA 839

RESULT 32
US-09-489-039A-11632
; Sequence 11632, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 11632
; LENGTH: 458
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-11632
Query Match 4.9%; Score 133.5; DB 2; Length 458;
Best Local Similarity 20.0%; Pred. No. 0.00024;

US-09-540-540-16500
; Sequence 16500, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 16500
; LENGTH: 272
; TYPE: PRT
; ORGANISM: Myxococcus xanthus
US-09-540-540-16500

Query Match 4.8%; Score 130.5; DB 2; Length 272;
Best Local Similarity 28.0%; Pred. No. 0.0002;
Matches 49; Conservative 22; Mismatches 73; Indels 31; Gaps 4;

Qy 299 SVOTRLISDTPAKGLDRDRKPPPIAGRLQDALKQPEKSVLVSPYFVPTKSGTDALAKLV 358
Db 10 TVLARGIADGDEDFETLRTV-----LLGALATARTSVRIPTYFLPDTALITALSVA 63

Qy 359 ODGIDVTVL-----TNSLOATDVAHVHSGVYKVKPGLLKAGIKLYELOPNHAVPATKDKGLT 415
Db 64 LRGVQVDILLPEKGNLPLVQWAAQLW-----QVLRPGCRIFLTAPPFD----- 108

Qy 416 GSSVTSLSHAKTFIVDGKRIFGSNLDRSARLNTMGVIVIESPKIAEQMERTLA 470
Db 109 -----HTKLMVVDGWSLIGSANWDRSURLNFEEVCEYDATALAMQEGVVA 156

RESULT 37
US-09-543-681A-5540
; Sequence 5540, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/09/543,681A
; CURRENT FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 5540
; LENGTH: 459
; TYPE: PRT
; ORGANISM: Proteus mirabilis
US-09-543-681A-5540

Query Match 4.8%; Score 130; DB 2; Length 459;
Best Local Similarity 20.7%; Pred. No. 0.00054;
Matches 98; Conservative 70; Mismatches 196; Indels 110; Gaps 21;

Qy 47 SRHNTSKPVLNDNIIQIRTPH-----NNGLSDIYLLDDPHEALAAAL---IESA 96
Db 1 TREIST---VNLSKLAKHQHQAQLPKLAQSVADVETL---YQTSAFRSTLLKYIAQA 54

Qy 97 EHSLDLQYIWRNDSIGRLFLNMYLA-----AERGVRVRLILDDNNT 139
Db 55 QKEIFITALVLEHDEAGEEILEALYAKQRPESLITVIWDHRAQRG-RIGVSAATNA 113

Qy 140 RGLDDLLALDSHPNIEVRLFN-PFVLKRWALGYLTFPRLNRRMHNKSFTADNRATIL 198
Db 114 ---DWYHKVAQHPDVPVIGYIPVNTRE--ALGVL-----HLKGFIFDD-TVIY 157

Qy 199 GGRNIGDEYFKVGEDTVFADLIDILATGVSVEYSHDFDRYWAHSAHNATR-----IIRSG 254
Db 158 SGASINNVYLH-----KLD-----KYRYDRYHIHNAELANTMKOFIVDSL 198

Qy 255 NIGKLGALGYNDSTSRHALLRYRETVEQSPLYQKIQTGRIDMOSVOTRILISDTPAKGLD 314
Db 199 LQSDAIQRL--DSEERVCAEIKNCIRQFRNFLRTHDGYDIKGNASNQLTVPVLGLG 255

Qy 315 RDRKPPIAGRLQDALKQPEKSVLVSPYFVPTKSGTDALAKLVQDGDIVTLVNSLOAT 374
Db 256 ---KCNILNKTISHLMAVTEKLVICTPYFNLPAILVRQISHLRNGKQVEIIGDKTAN 312

Qy 375 DVAAVHSGYVYKRP--LLKAGIKLYEL-----QSNHAVPATKDKGLTGSS 418
Db 313 DF-----YIPPEPFKIIGALPYLYEINLRKTFQRFQFIDNDQLTVRLWKGDNT--- 363

Qy 419 VTSLSHAKTFIVDGKRIFGSNLDRSARLNTMGVIVIESP--KIAEQMERTLA 470
Db 364 ---YHLKGMVWDDWQLITGNLNPRAWGLDLENAIILHDPHHELHEQRHKELA 414

RESULT 38
US-09-902-540-14817
; Sequence 14817, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 14817
; LENGTH: 155
; TYPE: PRT
; ORGANISM: Myxococcus xanthus
US-09-902-540-14817

Query Match 4.5%; Score 123; DB 2; Length 155;
Best Local Similarity 30.2%; Pred. No. 0.00044;
Matches 38; Conservative 17; Mismatches 63; Indels 8; Gaps 2;

Qy 78 LDDDPHEALAAARAAALIESAHSLSLDQYIWRNDSIGRLFLNMYLAAERGVRVRLILDDN 137
Db 38 VLESRRDAYPAMLEAIDAARTTILSLCSYIFDNDMAGRRFVEALSAAVKRGVEVRLVDAV 97

Qy 138 NTRGLDDLLALDSHPNIEVRLFNPFVLKRWALGYLTFPRLNRRMHNKSFTADNRATI 197
Db 98 GSRYTWPPIILGRLLRRAGVRAARFLP-SLMPYR-----LPFANLRNHRKLMVVDGVRGF 149

Qy 198 LGGRI 203
Db 150 TGGMI 155

RESULT 39
US-09-107-532A-5960
; Sequence 5960, Application US/09107532A
; Patent No. 6583275
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; NUMBER OF SEQUENCES: 7310
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: May 2, 2006, 06:24:35 ; Search time 166 Seconds
(without alignments)
1321.448 Million cell updates/sec

Title: US-10-665-990A-14

Perfect score: 2720

Sequence: 1 MHTDPKIQAMPSETISPMKT.....KLMKRIAAKILSLPIEGLL 525

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications AA Main:

- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
- 3: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
- 4: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
- 6: /cgn2_6/ptodata/1/pubpaa/US11_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2720	100.0	525	5	US-10-665-990A-14
2	2712	99.7	525	5	US-10-665-990A-16
3	2708	99.6	525	5	US-10-665-990A-18
4	2643	97.2	525	4	US-10-066-551-4
5	2643	97.2	525	5	US-10-665-990A-4
6	2643	97.2	525	5	US-10-665-990A-20
7	632	23.2	502	4	US-10-335-977-9085
8	632	23.2	502	4	US-10-335-977-9086
9	590	21.7	428	3	US-09-861-752A-356
10	359	13.2	224	4	US-10-335-977-9084
11	296	10.9	486	4	US-10-418-861B-59
12	291	10.7	495	4	US-10-724-972A-3854
13	286.5	10.5	493	3	US-09-828-523A-22
14	286.5	10.5	493	3	US-09-966-521-18
15	286.5	10.5	493	4	US-10-429-094-18
16	286.5	10.5	502	3	US-09-828-523A-90
17	286.5	10.5	502	3	US-09-966-521-84
18	286.5	10.5	502	4	US-10-429-094-84
19	274.5	10.1	206	4	US-10-335-977-9083
20	259.5	9.5	494	4	US-10-724-972A-4379
21	252	9.3	422	4	US-10-627-476-78
22	252	9.3	500	3	US-09-848-726-2
23	252	9.3	500	3	US-09-738-626-6514
24	226	8.3	404	4	US-10-156-761-8603
25	219	8.1	510	4	US-10-474-776-225
26	219	8.1	510	5	US-10-472-928-212
27	215	7.9	518	5	US-10-617-320-3177

28	205	7.5	363	5	US-10-501-282-2384	Sequence 2384, Ap
29	205	7.5	448	5	US-10-501-282-2386	Sequence 2386, Ap
30	205	7.5	481	5	US-10-501-282-2388	Sequence 2388, Ap
31	205	7.5	493	5	US-10-501-282-2390	Sequence 2390, Ap
32	205	7.5	512	5	US-10-501-282-2392	Sequence 2392, Ap
33	175	6.4	280	5	US-10-450-763-31511	Sequence 31511, A
34	175	6.4	1097	5	US-10-450-763-31510	Sequence 31510, A
35	168.5	6.2	525	5	US-10-474-792-354	Sequence 354, App
36	141.5	5.2	476	4	US-10-312-273-75	Sequence 75, Appl
37	133.5	4.9	932	4	US-10-137-129A-6	Sequence 6, Appl
38	127	4.7	474	5	US-10-498-327-79	Sequence 79, Appl
39	124.5	4.6	548	4	US-10-408-765A-2203	Sequence 2203, Ap
40	122.5	4.5	553	4	US-10-424-599-215690	Sequence 215690, A
41	121.5	4.5	1683	4	US-10-369-493-22273	Sequence 22273, A
42	116.5	4.3	908	4	US-10-032-585-7642	Sequence 7642, Ap
43	116	4.3	851	4	US-10-437-963-120957	Sequence 120957, A
44	115.5	4.2	1046	4	US-10-437-963-175603	Sequence 175603, A
45	113.5	4.2	494	6	US-11-097-143-2316	Sequence 2316, Ap

ALIGNMENTS

RESULT 1

US-10-665-990A-14
; Sequence 14, Application US/10665990A
; Publication No. US20040253222A1
; GENERAL INFORMATION:
; APPLICANT: Apicella, Michael A.
; APPLICANT: Edwards, Jennifer L.
; TITLE OF INVENTION: Vaccine and compositions for the prevention and treatment of Nei
; TITLE OF INVENTION: infections
; FILE REFERENCE: 17023-031001
; CURRENT APPLICATION NUMBER: US/10/665,990A
; PRIOR FILING DATE: 2003-09-19
; PRIOR APPLICATION NUMBER: US 10/621,184
; PRIOR FILING DATE: 2003-07-15
; PRIOR APPLICATION NUMBER: US 10/066,551
; PRIOR FILING DATE: 2002-01-31
; PRIOR APPLICATION NUMBER: US 60/344,452
; PRIOR FILING DATE: 2001-10-23
; PRIOR APPLICATION NUMBER: US 60/310,356
; PRIOR FILING DATE: 2001-08-06
; PRIOR APPLICATION NUMBER: US 60/266,070
; PRIOR FILING DATE: 2001-01-31
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14
; LENGTH: 525
; TYPE: PRT
; ORGANISM: Neisseria gonorrhoeae 1291
US-10-665-990A-14

Query Match	100.0%;	Score	2720;	DB	5;	Length	525;
Best Local Similarity	100.0%;	Pred. No.	3.5e-233;				
Matches	525;	Conservative	0;	Mismatches	0;	Indels	0;
						Gaps	0;
QY	1	MHTDPKIQAMPSETISPMKTRSLISLLCLLSCSSWLPPEERTESRHNTPKPVLLDN	60				
DB	1	MHTDPKIQAMPSETISPMKTRSLISLLCLLSCSSWLPPEERTESRHNTPKPVLLDN	60				
QY	61	ILQIRHTPHNGLSDIYLLDDPHEALAAPAAALIESAEHSLDLOYYIWRNDISGRLLFNLM	120				
DB	61	ILQIRHTPHNGLSDIYLLDDPHEALAAPAAALIESAEHSLDLOYYIWRNDISGRLLFNLM	120				
QY	121	YLAERGVRVRLLLDDNNTRGLDDLLALDSDHPNIEVLFNPFVLRKRWALGYLTDFPRL	180				
DB	121	YLAERGVRVRLLLDDNNTRGLDDLLALDSDHPNIEVLFNPFVLRKRWALGYLTDFPRL	180				
QY	181	NRRMANKSTADNRATILGGRNIGDEYFKVGEDTVFADLLDILATGSVGVSEVSHDFORYWA	240				
DB	181	NRRMANKSTADNRATILGGRNIGDEYFKVGEDTVFADLLDILATGSVGVSEVSHDFORYWA	240				

Qy 301 QTRLISDTPAKGLDRDRKPPKPIAGRLQDALKOPEKSVYLVSPYFVPTKSGTDALAKLVQD 360
Db 301 QTRLISDTPAKGLDRDRKPPKPIAGRLQDALKOPEKSVYLVSPYFVPTKSGTDALAKLVQD 360
Qy 361 GIDVTVLNTSLQATDVAHVSGYKVRKPKLLKAGIKLYELOPNHAVPATKDKGLTGSSVT 420
Db 361 GIDVTVLNTSLQATDVAHVSGYKVRKPKLLKAGIKLYELOPNHAVPATKDKGLTGSSVT 420
Qy 421 SLHAKTFIVDGKRIFIGSFNLDPRSARLNTMGVVIESPKEAQWERTLADTTPEYAYRV 480
Db 421 SLHAKTFIVDGKRIFIGSFNLDPRSARLNTMGVVIESPKEAQWERTLADTTPEYAYRV 480
Qy 481 TLDKHNRLQWHDPATRKTYPNEPEAKLWKRIAAKILSLPIEGLL 525
Db 481 TLDKHNRLQWHDPATRKTYPNEPEAKLWKRIAAKILSLPIEGLL 525

RESULT 4
US-10-066-551-4
; Sequence 4, Application US/10066551
; Publication No. US20030100071A1
; GENERAL INFORMATION:
; APPLICANT: Apicella, M. A.
; APPLICANT: Edwards, J. L.
; APPLICANT: Gibson, B. W.
; APPLICANT: Scheffler, K.
; APPLICANT: Brown, E.
; TITLE OF INVENTION: Vaccine and compositions for the
; prevention and treatment of Neisserial infections
; FILE REFERENCE: 875.045U1
; CURRENT APPLICATION NUMBER: US/10/066.551
; PRIOR FILING DATE: 2002-01-31
; PRIOR APPLICATION NUMBER: US 60/344,452
; PRIOR FILING DATE: 2001-10-23
; PRIOR APPLICATION NUMBER: US 60/310,356
; PRIOR FILING DATE: 2001-08-06
; PRIOR APPLICATION NUMBER: US 60/266,070
; PRIOR FILING DATE: 2001-01-31
; NUMBER OF SEQ ID NOS: 12
; SEQ ID NO 4
; LENGTH: 525
; TYPE: PRT
; ORGANISM: Neisseria gonorrhoeae
US-10-066-551-4

Query Match 97.2%; Score 2643; DB 4; Length 525;
Best Local Similarity 97.3%; Pred. No. 2.6e-226;
Matches 511; Conservative 5; Mismatches 9; Indels 0; Gaps 0;
Qy 1 MHTDPKIQAMPSETISPMKTRSLISLLCLLCCSSWLPLEERTESRHFNTSKPVLND 60
Db 1 MRANPKTQAMPSETISLMKTRSLISLLCLLCCSSWLPLEERTESRHFNTSKPVLND 60
Qy 61 ILQIRHTPHNGLSDIYLLDDPHEALAAARALIESAHSLDQYIWRNDISGRLLPNLM 120
Db 61 ILQIRHTPHNGLSDIYLLDDPHEAFAARAALIESAHSLDQYIWRNDISGRLLPNLV 120
Qy 121 YLAAEGRVRLLLDDNNTGRLDLDLALDHPNIEVRLNPFVLRKWRALGYLTDPPRL 180
Db 121 YLAAEGRVRLLLDDNNTGRLDLDLALDHPNIEVRLNPFVLRKWRALGYLTDPPRL 180
Qy 181 NRRMHKSFADNRATILGGRNIGDEYFKVGEDTVFADLDLILATGSVVGEVSHDFRYWA 240
Db 181 NRRMHKSFADNRATILGGRNIGDEYFKVGEDTVFADLDLILATGSVVGEVSHDFRYWA 240
Qy 241 SHSAHNATRIIRSGNIGKQALGYNDETSRHALLRYRETVESQPLKYQKIQTRIDWQSV 300
Db 241 SHSAHNATRIIRSGNIGKQALGYNDETSRHALLRYRETVESQPLKYQKIQTRIDWQSV 300
Qy 301 QTRLISDTPAKGLDRDRKPPKPIAGRLQDALKOPEKSVYLVSPYFVPTKSGTDALAKLVQD 360
Db 301 QTRLISDTPAKGLDRDRKPPKPIAGRLQDALKOPEKSVYLVSPYFVPTKSGTDALAKLVQD 360

Qy 361 GIDVTVLNTSLQATDVAHVSGYKVRKPKLLKAGIKLYELOPNHAVPATKDKGLTGSSVT 420
Db 361 GIDVTVLNTSLQATDVAHVSGYKVRKPKLLKAGIKLYELOPNHAVPATKDKGLTGSSVT 420
Qy 421 SLHAKTFIVDGKRIFIGSFNLDPRSARLNTMGVVIESPKEAQWERTLADTTPEYAYRV 480
Db 421 SLHAKTFIVDGKRIFIGSFNLDPRSARLNTMGVVIESPKEAQWERTLADTTPEYAYRV 480
Qy 481 TLDKHNRLQWHDPATRKTYPNEPEAKLWKRIAAKILSLPIEGLL 525
Db 481 TLDKHNRLQWHDPATRKTYPNEPEAKLWKRIAAKILSLPIEGLL 525

RESULT 5
US-10-665-990A-4
; Sequence 4, Application US/10665990A
; Publication No. US20040253222A1
; GENERAL INFORMATION:
; APPLICANT: Apicella, Michael A.
; APPLICANT: Edwards, Jennifer L.
; TITLE OF INVENTION: Vaccine and compositions for the prevention and treatment of Nei
; TITLE OF INVENTION: infections
; FILE REFERENCE: 17023-031001
; CURRENT APPLICATION NUMBER: US/10/665,990A
; PRIOR FILING DATE: 2003-09-19
; PRIOR APPLICATION NUMBER: US 10/621,184
; PRIOR FILING DATE: 2003-07-15
; PRIOR APPLICATION NUMBER: US 10/066,551
; PRIOR FILING DATE: 2002-01-31
; PRIOR APPLICATION NUMBER: US 60/344,452
; PRIOR FILING DATE: 2001-10-23
; PRIOR APPLICATION NUMBER: US 60/310,356
; PRIOR FILING DATE: 2001-08-06
; PRIOR APPLICATION NUMBER: US 60/266,070
; PRIOR FILING DATE: 2001-01-31
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 525
; TYPE: PRT
; ORGANISM: Neisseria gonorrhoeae
US-10-665-990A-4

Query Match 97.2%; Score 2643; DB 5; Length 525;
Best Local Similarity 97.3%; Pred. No. 2.6e-226;
Matches 511; Conservative 5; Mismatches 9; Indels 0; Gaps 0;
Qy 1 MHTDPKIQAMPSETISPMKTRSLISLLCLLCCSSWLPLEERTESRHFNTSKPVLND 60
Db 1 MRANPKTQAMPSETISLMKTRSLISLLCLLCCSSWLPLEERTESRHFNTSKPVLND 60
Qy 61 ILQIRHTPHNGLSDIYLLDDPHEALAAARALIESAHSLDQYIWRNDISGRLLPNLM 120
Db 61 ILQIRHTPHNGLSDIYLLDDPHEAFAARAALIESAHSLDQYIWRNDISGRLLPNLV 120
Qy 121 YLAAEGRVRLLLDDNNTGRLDLDLALDHPNIEVRLNPFVLRKWRALGYLTDPPRL 180
Db 121 YLAAEGRVRLLLDDNNTGRLDLDLALDHPNIEVRLNPFVLRKWRALGYLTDPPRL 180
Qy 181 NRRMHKSFADNRATILGGRNIGDEYFKVGEDTVFADLDLILATGSVVGEVSHDFRYWA 240
Db 181 NRRMHKSFADNRATILGGRNIGDEYFKVGEDTVFADLDLILATGSVVGEVSHDFRYWA 240
Qy 241 SHSAHNATRIIRSGNIGKQALGYNDETSRHALLRYRETVESQPLKYQKIQTRIDWQSV 300
Db 241 SHSAHNATRIIRSGNIGKQALGYNDETSRHALLRYRETVESQPLKYQKIQTRIDWQSV 300
Qy 301 QTRLISDTPAKGLDRDRKPPKPIAGRLQDALKOPEKSVYLVSPYFVPTKSGTDALAKLVQD 360
Db 301 QTRLISDTPAKGLDRDRKPPKPIAGRLQDALKOPEKSVYLVSPYFVPTKSGTDALAKLVQD 360
Qy 361 GIDVTVLNTSLQATDVAHVSGYKVRKPKLLKAGIKLYELOPNHAVPATKDKGLTGSSVT 420
Db 361 GIDVTVLNTSLQATDVAHVSGYKVRKPKLLKAGIKLYELOPNHAVPATKDKGLTGSSVT 420

Db 361 GIDVTVLNSLOATDVAHVHGVYKVRKPLKAGIKLYELQPNHVAVPATKDKGLTGSSVT 420
Qy 421 SLHAKTFIVDGKRIFIGSFNLDPRSARLNTMGVVIESPKEAQEMERTLADTTPEYAYRV 480
Db 421 SLHAKTFIVDGKRIFIGSFNLDPRSARLNTMGVVIESPKEAQEMERTLADTSPEYAYRV 480
Qy 481 TLDKHNRLQWHDHPATRKTYPNPEPEAKLWKRIAANKILSLPIEGLL 525
Db 481 TLDKHNRLQWHDHPATRKTYPNPEPEAKLWKRIAANKILSLPIESLL 525

RESULT 6
US-10-665-990A-20
; Sequence 20, Application US/10665990A
; Publication No. US20040253222A1
; GENERAL INFORMATION:
; APPLICANT: Apicella, Michael A.
; APPLICANT: Edwards, Jennifer L.
; TITLE OF INVENTION: Vaccine and compositions for the prevention and treatment of Neis
; TITLE OF INVENTION: Infections
; FILE REFERENCE: 17023-031001
; CURRENT APPLICATION NUMBER: US/10/665,990A
; PRIOR FILING DATE: 2003-09-19
; PRIOR APPLICATION NUMBER: US 10/621,184
; PRIOR FILING DATE: 2003-07-15
; PRIOR APPLICATION NUMBER: US 10/066,551
; PRIOR FILING DATE: 2002-01-31
; PRIOR APPLICATION NUMBER: US 60/344,452
; PRIOR FILING DATE: 2001-10-23
; PRIOR APPLICATION NUMBER: US 60/310,356
; PRIOR FILING DATE: 2001-08-06
; PRIOR APPLICATION NUMBER: US 60/266,070
; PRIOR FILING DATE: 2001-01-31
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 20
; TYPE: PRT
; ORGANISM: Neisseria meningitidis B
US-10-665-990A-20

Query Match 97.2%; Score 2643; DB 5; Length 525;
Best Local Similarity 97.3%; Pred. No. 2.6e-226; Mismatches 5; Indels 0; Gaps 0;
Matches 511; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

Qy 1 MHTDQKAMPSETISPMKTRSLISLLCLLLCSCSSWLPPLLEERTESRHFNTSKPVLND 60
Db 1 MRANPKQAMPSETISLTKTRSLISLLCLLLCSCSSWLPPLLEERTESRHFNTSKPVLND 60

Qy 61 ILQIRHTPHNGLSDIYLLDDPHEALAAARALIESAEHSLDLQYIWRNDISGRLLFNLM 120
Db 61 ILQIRHTPHNGLSDIYLLNDPHEAFAARAALIESAEHSLDLQYIWRNDISGRLLFNLV 120

Qy 121 YLAERGVVRVLLDDNTRGLDILLALDHPNIEVLEFPVLRKWRALGYLTFPRL 180
Db 121 YLAERGVVRVLLDDNTRGLDILLALDHPNIEVLEFPVLRKWRALGYLTFPRL 180

Qy 181 NRRMKNKFTADNRATILGGRNIGDEYFKVGEDTVFADLDILATGSVVGEVSHDFRYWA 240
Db 181 NRRMKNKFTADNRATILGGRNIGDEYFKVGEDTVFADLDILATGSVVGEVSHDFRYWA 240

Qy 241 SHSAHNATRIIRSGNIGKGLQALQYNDETSRHALLRYRETVEQSPLYOKIQTGRIDWQSV 300
Db 241 SHSAHNATRIIRSGNIGKGLQALQYNDETSRHALLRYRETVEQSPLYOKIQTGRIDWQSV 300

Qy 301 QTRLISDTPAKGLDRDRKPPITAGRLQDALKOPEKSVVLVSPYFPTKSGTDLAKLVQD 360
Db 301 QTRLISDTPAKGLDRDRKPPITAGRLQDALKOPEKSVVLVSPYFPTKSGTDLAKLVQD 360

Qy 361 GIDVTVLNSLOATDVAHVHGVYKVRKPLKAGIKLYELQPNHVAVPATKDKGLTGSSVT 420
Db 361 GIDVTVLNSLOATDVAHVHGVYKVRKPLKAGIKLYELQPNHVAVPATKDKGLTGSSVT 420

Qy 421 SLHAKTFIVDGKRIFIGSFNLDPRSARLNTMGVVIESPKEAQEMERTLADTTPEYAYRV 480
Db 421 SLHAKTFIVDGKRIFIGSFNLDPRSARLNTMGVVIESPKEAQEMERTLADTSPEYAYRV 480
Qy 481 TLDKHNRLQWHDHPATRKTYPNPEPEAKLWKRIAANKILSLPIEGLL 525
Db 481 TLDKHNRLQWHDHPATRKTYPNPEPEAKLWKRIAANKILSLPIESLL 525

RESULT 7
US-10-335-977-9085
; Sequence 9085, Application US/10335977
; Publication No. US20040052799A1
; GENERAL INFORMATION:
; APPLICANT: DOUGLAS SMITH et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES
; RELATING TO HELICOBACTER PYLORI FOR
; DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 10031
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 28 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: Windows NT 4.0
; SOFTWARE: UNIX
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/335,977
; FILING DATE: 30-Dec-2002
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/993,002
; FILING DATE: 17-DEC-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Mandragouras, Amy E.
; REGISTRATION NUMBER: 36,207
; REFERENCE/DOCKET NUMBER: GTN-018
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)742-4214
; INFORMATION FOR SEQ ID NO: 9085:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 502 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHEICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: Helicobacter pylori
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (B) LOCATION 1...502
; SEQUENCE DESCRIPTION: SEQ ID NO: 9085:
US-10-335-977-9085

Query Match 23.2%; Score 632; DB 4; Length 502;
Best Local Similarity 32.8%; Pred. No. 4e-47;
Matches 161; Conservative 93; Mismatches 181; Indels 56; Gaps 17;

Qy 64 IRHTPHNGLSDIY-----LDDPHEALAAARALIESAEHSLDLQYIWRN 109
Db 30 ISYDPTTTIGSLYAKNLKENPKHSAAILLEDGFDALLHRVGLIRMSQKSIDMQYIYKN 89

Qy 110 DISGELLENMYLAERGVVRVLLDDNNTGLD----DLLALDSDHNIETVRLNPFVL 165
Db 90 DLSQVIAKELNLAANRGVVRILLDDN---GLOSDFSDIML-LNFHKNIEVKIENFYI 145

Qy 166 RKWRALGY---LTDFPRLNRRMKNKFTADNRATILGGRNIGDEYFKVGEDTVFADLDIL 222

Db 146 RN-KGLRYFEMLADYERIKRMHKLFIVDNFAVIGGRNIGDNYFNDLDTNFDLIDLAL 204
Qy 223 ATGSGVGVSHDFRYWASHAHNATRIIRS-----GNIGKGLQALGVNDETSRHAL 274
Db 205 PFGGVASAKAKESFENYRFRHSIPVS--LLRTHKRLKNNVKEIAKLHEKIPISAEADANEFE 263
Qy 275 LRYRETVEQSPLYQ-KIQTGRIDWQSVOTRLISDTPAKGLDRDRRKPPIAGRLQDALKQP 333
Db 264 KKVNDFIERFQKYQVPIYG-----NAIFLADLPK-IDTPLYS-PIKIAFEKALKNA 314
Qy 334 EKSIVLVSPYFVPTKSGTDALAKLVQDIDVTVLTNLSQATDVAHVSGYVYKRPKLKA 393
Db 315 KDSVFIASSYFIPGKIMKIFKNQISKGIELNLTNLSSTDAIVVYGAWERYNKLVRM 374
Qy 394 GIKLYELOPNHAPVATKDKGLTSGSVTSLHAKTIFVDGKRIFIGSFNLDPRSARLNTMG 453
Db 375 GANVYEIRNDFNRQIKGR----FTKISLHGKTVFDDALTLLGSFNIDPRSAINTESA 431
Qy 454 VVIESPKIAEQMERTLADTTPYAYRVTLDKHNLQWHDHPATRKTY----PNEPEAKLWK 509
Db 432 VLFNPNFSAKRVRLSLKDHA-QQSWHLVLYRH-RVIME--ATEBGILIHKNSPDTSFPL 487
Qy 510 RIAAKTILSLP 520
Db 488 RLKIEWSKVLP 498
RESULT 8
US-10-335-977-9086
; Sequence 9086, Application US/10335977
; Publication No. US20040052799A1
; GENERAL INFORMATION:
; APPLICANT: DOUGLAS SMITH et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES
; RELATING TO HELICOBACTER PYLORI FOR
; DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 10031
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 28 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: Windows NT 4.0
; SOFTWARE: UNIX
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/335,977
; FILING DATE: 30-Dec-2002
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/993,002
; FILING DATE: 17-DEC-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Mandragouras, Amy E.
; REGISTRATION NUMBER: 36,207
; REFERENCE/DOCKET NUMBER: GTN-018
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)742-4214
; INFORMATION FOR SEQ ID NO: 9086:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 502 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: Helicobacter pylori
; FEATURE:
; NAME/KEY: misc_feature

; LOCATION: (B) LOCATION 1...502
; SEQUENCE DESCRIPTION: SEQ ID NO: 9086:
US-10-335-977-9086
Query Match 23.2%; Score 632; DB 4; Length 502;
Best Local Similarity 32.8%; Pred. No. 4e-47;
Matches 161; Conservative 93; Mismatches 181; Indels 56; Gaps 17;
Qy 64 IRHTPHNNGSLDIY-----LLDDPHEALAAARALIESAEHSLDLOYIWRN 109
Db 30 ISVDYPTTTTIGSYAKNKLKPKHSAAILLEDGFDALLHRVGLIRMSQKSIDMQTYTKN 89
Qy 110 DISGRILFNLMYLAERGVRLLDNNTRGLD---DLLLALDSHPNLEVRLEFNPVL 165
Db 90 DLSQVIKELLNANRGVVRILLDDN---GLDSFSDIML-LNFHKNIEVKIFNPYYI 145
Qy 166 RKWRALGY---LTFPRLNRRMKNKSTADNRATILGGRNIGDEYFKVGEDTVFADLDIL 222
Db 146 RN-KGLRYFEMLADYERIKRMHKLFIVDNFAVIGGRNIGDNYFNDLDTNFDLIDLAL 204
Qy 223 ATGSGVGVSHDFRYWASHAHNATRIIRS-----GNIGKGLQALGVNDETSRHAL 274
Db 205 PFGGVASAKAKESFENYRFRHSIPVS--LLRTHKRLKNNVKEIAKLHEKIPISAEADANEFE 263
Qy 275 LRYRETVEQSPLYQ-KIQTGRIDWQSVOTRLISDTPAKGLDRDRRKPPIAGRLQDALKQP 333
Db 264 KKVNDFIERFQKYQVPIYG-----NAIFLADLPK-IDTPLYS-PIKIAFEKALKNA 314
Qy 334 EKSIVLVSPYFVPTKSGTDALAKLVQDIDVTVLTNLSQATDVAHVSGYVYKRPKLKA 393
Db 315 KDSVFIASSYFIPGKIMKIFKNQISKGIELNLTNLSSTDAIVVYGAWERYNKLVRM 374
Qy 394 GIKLYELOPNHAPVATKDKGLTSGSVTSLHAKTIFVDGKRIFIGSFNLDPRSARLNTMG 453
Db 375 GANVYEIRNDFNRQIKGR----FTKISLHGKTVFDDALTLLGSFNIDPRSAINTESA 431
Qy 454 VVIESPKIAEQMERTLADTTPYAYRVTLDKHNLQWHDHPATRKTY----PNEPEAKLWK 509
Db 432 VLFNPNFSAKRVRLSLKDHA-QQSWHLVLYRH-RVIME--ATEBGILIHKNSPDTSFPL 487
Qy 510 RIAAKTILSLP 520
Db 488 RLKIEWSKVLP 498
RESULT 9
US-09-881-752A-356
; Sequence 356, Application US/09881752A
; Patent No. US20020115078A1
; GENERAL INFORMATION:
; APPLICANT: Kleanthous, Harold
; APPLICANT: Al-Garawi, Amal
; APPLICANT: Miller, Charles
; APPLICANT: Tomb, Jean-Francois
; APPLICANT: Oomen, Raymond P.
; TITLE OF INVENTION: Identification of Polynucleotides
; TITLE OF INVENTION: Encoding No. US20020115078A1e1 Helicobacter Polypeptides in the
; FILE REFERENCE: 06132/041002
; CURRENT APPLICATION NUMBER: US/09/881,752A
; CURRENT FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: US 08/833,457
; PRIOR FILING DATE: 1997-04-01
; NUMBER OF SEQ ID NOS: 370
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 356
; LENGTH: 428
; TYPE: PRT
; ORGANISM: Helicobacter pylori
US-09-881-752A-356
Query Match 21.7%; Score 590; DB 3; Length 428;
Best Local Similarity 32.3%; Pred. No. 1.8e-43;


```
Qy 376 VAAHSGVVKYKPKLLKAGIKLYELQPNHAPATKDKLTGSSVTSLSHAKTFIVDGKRF 435
Db 381 HPLVYATFSNASDLSGSKIY-----TVENG-----IHSKMLIDDEIVS 423
Qy 436 IGSFNLDPRSARLNTMGVWIESPKIABQMERTLTADTTPEYAYRVTLDKHNLQWHPAT 495
Db 424 VGTANMDFRSFELNFVNAFYVDENLAKDL-----RVAYEHDTSKQL-----T 468
Qy 496 RKTYPNEPEAKLWKRIAAKILS 517
Db 469 KESYANRPLSVKFKESLAKLVS 490

RESULT 14
US-09-966-521-18
; Sequence 18, Application US/09966521
; Publication No. US20030087321A1
; GENERAL INFORMATION:
; APPLICANT: TOMICH, Che-Shen
; APPLICANT: QUINN, Cheryl
; APPLICANT: ARVIDSON, Staffan
; APPLICANT: HARRIS, Douglas
; APPLICANT: MOTT, John
; TITLE OF INVENTION: ANTIMICROBIAL METHODS AND MATERIALS
; CURRENT APPLICATION NUMBER: US/09/966,521
; CURRENT FILING DATE: 2001-09-28
; NUMBER OF SEQ ID NOS: 145
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 18
; LENGTH: 493
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-966-521-18

Query Match 10.5%; Score 286.5; DB 3; Length 493;
Best Local Similarity 21.1%; Pred. No. 2.4e-16;
Matches 106; Conservative 86; Mismatches 169; Indels 141; Gaps 16;

Qy 42 EERTESRHNTSKPVLLD-----NILQIRHTPHNGLSDIYLLDDPHEALAAARALIES 95
Db 104 DNKQVQHDLVRMLMDQDGLTENNKVDHFIDGNLDYQVLKD-----IKN 151
Qy 96 AEHSLDLOYITWRNDISGRLLFNMLYLAABGVVRVRLDDNNTRGLDLDLALLDLSHPI 155
Db 152 AKEYIHLEYTYTALDGLGRILHLEEKLGKQLEVKIYDDVGSK-----NV 198
Qy 156 EVRLFNPFLVKRWALG-----YLTDFP-----RLNRRMHNKSFADNRATILGGRNIGD 205
Db 199 KMANFDHF-----KSLGGEVAFASKLPLNFRMNNHRKIIVIDQGLGVGGFNIGD 253
Qy 206 EYFKVGEDTVFADLDILATGVSVEVSHDPRYVASHSAHNATRII-----RSGNIGK 258
Db 254 EYLGKGLGYWRDTHLRIGQDAVDALQRLRFDLWNS-QAHRPQFEYDVKYPKNGPLG- 311
Qy 259 GLQALGYNDETSRHALLRYRETVESQPLYQKIOTGRIDWQSVQ---TRLISDTTAKGLDR 315
Db 312 -----NSPIQIAASGPASDWHQIEYGYTKMWSA----- 340
Qy 316 DRRKPPITAGRLQDALKQPEKSVYLVSPYFVPTKSGTDALAKLVODGIDVTVLTNSLOATD 375
Db 341 -----KKSVMYLOSQPYFIPDINSYINAIKTAAGSGVDVHLMIPC--KPD 380
Qy 376 VAAHSGVVKYKPKLLKAGIKLYELQPNHAPATKDKLTGSSVTSLSHAKTFIVDGKRF 435
Db 381 HPLVYATFSNASDLSGSKIY-----TVENG-----IHSKMLIDDEIVS 423
Qy 436 IGSFNLDPRSARLNTMGVWIESPKIABQMERTLTADTTPEYAYRVTLDKHNLQWHPAT 495
Db 424 VGTANMDFRSFELNFVNAFYVDENLAKDL-----RVAYEHDTSKQL-----T 468
Qy 496 RKTYPNEPEAKLWKRIAAKILS 517
Db 469 KESYANRPLSVKFKESLAKLVS 490

RESULT 15
US-10-429-094-18
; Sequence 18, Application US/10429094
; Publication No. US20030180821A1
; GENERAL INFORMATION:
; APPLICANT: TOMICH, Che-Shen
; APPLICANT: QUINN, Cheryl
; APPLICANT: ARVIDSON, Staffan
; APPLICANT: HARRIS, Douglas
; APPLICANT: MOTT, John
; TITLE OF INVENTION: ANTIMICROBIAL METHODS AND MATERIALS
; CURRENT APPLICATION NUMBER: US/10/429,094
; CURRENT FILING DATE: 2003-05-02
; PRIOR APPLICATION NUMBER: US 09/966,521
; PRIOR FILING DATE: 2001-09-28
; NUMBER OF SEQ ID NOS: 145
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 18
; LENGTH: 493
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-10-429-094-18

Query Match 10.5%; Score 286.5; DB 4; Length 493;
Best Local Similarity 21.1%; Pred. No. 2.4e-16;
Matches 106; Conservative 86; Mismatches 169; Indels 141; Gaps 16;

Qy 42 EERTESRHNTSKPVLLD-----NILQIRHTPHNGLSDIYLLDDPHEALAAARALIES 95
Db 104 DNKQVQHDLVRMLMDQDGLTENNKVDHFIDGNLDYQVLKD-----IKN 151
Qy 96 AEHSLDQYIYITWRNDISGRLLFNMLYLAABGVVRVRLDDNNTRGLDLDLALLDLSHPI 155
Db 152 AKEYIHLEYTYTALDGLGRILHLEEKLGKQLEVKIYDDVGSK-----NV 198
Qy 156 EVRLFNPFLVKRWALG-----YLTDFP-----RLNRRMHNKSFADNRATILGGRNIGD 205
Db 199 KMANFDHF-----KSLGGEVAFASKLPLNFRMNNHRKIIVIDQGLGVGGFNIGD 253
Qy 206 EYFKVGEDTVFADLDILATGVSVEVSHDPRYVASHSAHNATRII-----RSGNIGK 258
Db 254 EYLGKGLGYWRDTHLRIGQDAVDALQRLRFDLWNS-QAHRPQFEYDVKYPKNGPLG- 311
Qy 259 GLQALGYNDETSRHALLRYRETVESQPLYQKIOTGRIDWQSVQ---TRLISDTTAKGLDR 315
Db 312 -----NSPIQIAASGPASDWHQIEYGYTKMWSA----- 340
Qy 316 DRRKPPITAGRLQDALKQPEKSVYLVSPYFVPTKSGTDALAKLVODGIDVTVLTNSLOATD 375
Db 341 -----KKSVMYLOSQPYFIPDINSYINAIKTAAGSGVDVHLMIPC--KPD 380
Qy 376 VAAHSGVVKYKPKLLKAGIKLYELQPNHAPATKDKLTGSSVTSLSHAKTFIVDGKRF 435
Db 381 HPLVYATFSNASDLSGSKIY-----TVENG-----IHSKMLIDDEIVS 423
Qy 436 IGSFNLDPRSARLNTMGVWIESPKIABQMERTLTADTTPEYAYRVTLDKHNLQWHPAT 495
Db 424 VGTANMDFRSFELNFVNAFYVDENLAKDL-----RVAYEHDTSKQL-----T 468
Qy 496 RKTYPNEPEAKLWKRIAAKILS 517
Db 469 KESYANRPLSVKFKESLAKLVS 490

RESULT 16
US-09-828-523A-90
; Sequence 90, Application US/09828523A
; Patent No. US20020168697A1
; GENERAL INFORMATION:
```

```
Qy 376 VAAHSGVVKYKPKLLKAGIKLYELQPNHAPATKDKLTGSSVTSLSHAKTFIVDGKRF 435
Db 381 HPLVYATFSNASDLSGSKIY-----TVENG-----IHSKMLIDDEIVS 423
Qy 436 IGSFNLDPRSARLNTMGVWIESPKIABQMERTLTADTTPEYAYRVTLDKHNLQWHPAT 495
Db 424 VGTANMDFRSFELNFVNAFYVDENLAKDL-----RVAYEHDTSKQL-----T 468
Qy 496 RKTYPNEPEAKLWKRIAAKILS 517
Db 469 KESYANRPLSVKFKESLAKLVS 490

RESULT 14
US-09-966-521-18
; Sequence 18, Application US/09966521
; Publication No. US20030087321A1
; GENERAL INFORMATION:
; APPLICANT: TOMICH, Che-Shen
; APPLICANT: QUINN, Cheryl
; APPLICANT: ARVIDSON, Staffan
; APPLICANT: HARRIS, Douglas
; APPLICANT: MOTT, John
; TITLE OF INVENTION: ANTIMICROBIAL METHODS AND MATERIALS
; CURRENT APPLICATION NUMBER: US/09/966,521
; CURRENT FILING DATE: 2001-09-28
; NUMBER OF SEQ ID NOS: 145
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 18
; LENGTH: 493
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-966-521-18

Query Match 10.5%; Score 286.5; DB 3; Length 493;
Best Local Similarity 21.1%; Pred. No. 2.4e-16;
Matches 106; Conservative 86; Mismatches 169; Indels 141; Gaps 16;

Qy 42 EERTESRHNTSKPVLLD-----NILQIRHTPHNGLSDIYLLDDPHEALAAARALIES 95
Db 104 DNKQVQHDLVRMLMDQDGLTENNKVDHFIDGNLDYQVLKD-----IKN 151
Qy 96 AEHSLDLOYITWRNDISGRLLFNMLYLAABGVVRVRLDDNNTRGLDLDLALLDLSHPI 155
Db 152 AKEYIHLEYTYTALDGLGRILHLEEKLGKQLEVKIYDDVGSK-----NV 198
Qy 156 EVRLFNPFLVKRWALG-----YLTDFP-----RLNRRMHNKSFADNRATILGGRNIGD 205
Db 199 KMANFDHF-----KSLGGEVAFASKLPLNFRMNNHRKIIVIDQGLGVGGFNIGD 253
Qy 206 EYFKVGEDTVFADLDILATGVSVEVSHDPRYVASHSAHNATRII-----RSGNIGK 258
Db 254 EYLGKGLGYWRDTHLRIGQDAVDALQRLRFDLWNS-QAHRPQFEYDVKYPKNGPLG- 311
Qy 259 GLQALGYNDETSRHALLRYRETVESQPLYQKIOTGRIDWQSVQ---TRLISDTTAKGLDR 315
Db 312 -----NSPIQIAASGPASDWHQIEYGYTKMWSA----- 340
Qy 316 DRRKPPITAGRLQDALKQPEKSVYLVSPYFVPTKSGTDALAKLVODGIDVTVLTNSLOATD 375
Db 341 -----KKSVMYLOSQPYFIPDINSYINAIKTAAGSGVDVHLMIPC--KPD 380
Qy 376 VAAHSGVVKYKPKLLKAGIKLYELQPNHAPATKDKLTGSSVTSLSHAKTFIVDGKRF 435
Db 381 HPLVYATFSNASDLSGSKIY-----TVENG-----IHSKMLIDDEIVS 423
Qy 436 IGSFNLDPRSARLNTMGVWIESPKIABQMERTLTADTTPEYAYRVTLDKHNLQWHPAT 495
Db 424 VGTANMDFRSFELNFVNAFYVDENLAKDL-----RVAYEHDTSKQL-----T 468
Qy 496 RKTYPNEPEAKLWKRIAAKILS 517
Db 469 KESYANRPLSVKFKESLAKLVS 490

RESULT 15
US-10-429-094-18
; Sequence 18, Application US/10429094
; Publication No. US20030180821A1
; GENERAL INFORMATION:
; APPLICANT: TOMICH, Che-Shen
; APPLICANT: QUINN, Cheryl
; APPLICANT: ARVIDSON, Staffan
; APPLICANT: HARRIS, Douglas
; APPLICANT: MOTT, John
; TITLE OF INVENTION: ANTIMICROBIAL METHODS AND MATERIALS
; CURRENT APPLICATION NUMBER: US/10/429,094
; CURRENT FILING DATE: 2003-05-02
; PRIOR APPLICATION NUMBER: US 09/966,521
; PRIOR FILING DATE: 2001-09-28
; NUMBER OF SEQ ID NOS: 145
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 18
; LENGTH: 493
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-10-429-094-18

Query Match 10.5%; Score 286.5; DB 4; Length 493;
Best Local Similarity 21.1%; Pred. No. 2.4e-16;
Matches 106; Conservative 86; Mismatches 169; Indels 141; Gaps 16;

Qy 42 EERTESRHNTSKPVLLD-----NILQIRHTPHNGLSDIYLLDDPHEALAAARALIES 95
Db 104 DNKQVQHDLVRMLMDQDGLTENNKVDHFIDGNLDYQVLKD-----IKN 151
Qy 96 AEHSLDQYIYITWRNDISGRLLFNMLYLAABGVVRVRLDDNNTRGLDLDLALLDLSHPI 155
Db 152 AKEYIHLEYTYTALDGLGRILHLEEKLGKQLEVKIYDDVGSK-----NV 198
Qy 156 EVRLFNPFLVKRWALG-----YLTDFP-----RLNRRMHNKSFADNRATILGGRNIGD 205
Db 199 KMANFDHF-----KSLGGEVAFASKLPLNFRMNNHRKIIVIDQGLGVGGFNIGD 253
Qy 206 EYFKVGEDTVFADLDILATGVSVEVSHDPRYVASHSAHNATRII-----RSGNIGK 258
Db 254 EYLGKGLGYWRDTHLRIGQDAVDALQRLRFDLWNS-QAHRPQFEYDVKYPKNGPLG- 311
Qy 259 GLQALGYNDETSRHALLRYRETVESQPLYQKIOTGRIDWQSVQ---TRLISDTTAKGLDR 315
Db 312 -----NSPIQIAASGPASDWHQIEYGYTKMWSA----- 340
Qy 316 DRRKPPITAGRLQDALKQPEKSVYLVSPYFVPTKSGTDALAKLVODGIDVTVLTNSLOATD 375
Db 341 -----KKSVMYLOSQPYFIPDINSYINAIKTAAGSGVDVHLMIPC--KPD 380
Qy 376 VAAHSGVVKYKPKLLKAGIKLYELQPNHAPATKDKLTGSSVTSLSHAKTFIVDGKRF 435
Db 381 HPLVYATFSNASDLSGSKIY-----TVENG-----IHSKMLIDDEIVS 423
Qy 436 IGSFNLDPRSARLNTMGVWIESPKIABQMERTLTADTTPEYAYRVTLDKHNLQWHPAT 495
Db 424 VGTANMDFRSFELNFVNAFYVDENLAKDL-----RVAYEHDTSKQL-----T 468
Qy 496 RKTYPNEPEAKLWKRIAAKILS 517
Db 469 KESYANRPLSVKFKESLAKLVS 490

RESULT 16
US-09-828-523A-90
; Sequence 90, Application US/09828523A
; Patent No. US20020168697A1
; GENERAL INFORMATION:
```

; APPLICANT: The Pharmacia & Upjohn Company
; TITLE OF INVENTION: ANTIMICROBIAL METHODS AND MATERIALS
; FILE REFERENCE: 268.62120101
; CURRENT APPLICATION NUMBER: US/09/828,523A
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: 60/266,327
; PRIOR FILING DATE: 2000-04-06
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 90
; LENGTH: 502
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Amino acid sequence encoded by S. aureus coding region cloned for
; OTHER INFORMATION: expression in E. coli.
US-09-828-523A-90

Query Match 10.5%; Score 286.5; DB 3; Length 502;
Best Local Similarity 21.1%; Pred. No. 2.5e-16;
Matches 106; Conservative 86; Mismatches 169; Indels 141; Gaps 16;
Qy 42 EERTESRHFTSKPVLLD-----NILQIRHTPHNGLSDIYLLDDPHEALAAALIES 95
Db 105 DNKOVQKHDLVRMLMDQDGLTENKVDHFDIGNDLYDQVLKD-----IKN 152
Qy 96 AEHSLDLYQYIWRNDISGRLLFNLMYLAERGVRVRLLDNNTNRGLDLDLALDHPNI 155
Db 153 AKEVIHLEYTFALDGLGKILHALEKLGKLEKLVKLYDDVGSK-----NV 199
Qy 156 EVRLNFPVLRKWRALG-----YLTDFP-----RLNRRMNKSFADNRATILGGRNIGD 205
Db 200 KMANFDHF-----KSLGGEVEAFASKPLPLNFRMNRNRKIIIVIDGQLGVYGVGNIGD 254
Qy 206 EYFKVGEDTVFADLDILATGVSVEVSHDFDRYWAHSAHNATRII-----RSGNIGK 258
Db 255 EYLGKLGKLYWRDTHRIQGDALQRLFDLWNS--QAHRPQFEYDVVKYFPKNGPLG- 312
Qy 259 GLQALGYNDTSRHALLRYRETVEQSPLYQKIOTGRIDWQSVQ---TRLISDTPAKGLDR 315
Db 313 -----NSPIQTAASGPASDWHQIEGYTKMMSA----- 341
Qy 316 DRKPPPIAGRLQDALQKPEKSVLYSPYFPTKSGTDALAKLVODGIDVTVLNLSQATD 375
Db 342 -----KKSIVLQSPYFIPDINSYINAIKIAKSGVDVHLMIPC--KPD 381
Qy 376 VAAVHSGYKVRKPLKAGIKLYELQPNHAVPATKDKGLTGSSVTSLSHAKTIFVDGKRIF 435
Db 382 HPLVYWATFNSADLLSSGVKIY-----TYENG-----IHSKMLIDDEIVS 424
Qy 436 IGSFNLDPRGARLNTGEMVWVIESPKIAEQMERTLADTTPVAYRVTLDKHNRQWHPAT 495
Db 425 VGTANMDFRSEFELNFEVNAFVDENLAKDL-----RVAYEHDITKSKQL-----T 469
Qy 496 RKTYPNEPEAKLWKRIAAKILS 517
Db 470 KESYANRPLSVKPKESLAKLVS 491

RESULT 17
US-09-966-521-84
; Sequence 84, Application US/09966521
; Publication No. US20030087321A1
; GENERAL INFORMATION:
; APPLICANT: TOMICH, Che-Shen
; APPLICANT: QUINN, Cheryl
; APPLICANT: ARVIDSON, Staffan
; APPLICANT: HARRIS, Douglas
; APPLICANT: MOTT, John
; TITLE OF INVENTION: ANTIMICROBIAL METHODS AND MATERIALS
; FILE REFERENCE: 6212.N2
; CURRENT APPLICATION NUMBER: US/09/966,521
; CURRENT FILING DATE: 2001-09-28

; NUMBER OF SEQ ID NOS: 145
; SEQ ID NO 84
; LENGTH: 502
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Amino acid sequence encoded by S. aureus coding region cloned for
; OTHER INFORMATION: expression in E. coli
US-09-966-521-84

Query Match 10.5%; Score 286.5; DB 3; Length 502;
Best Local Similarity 21.1%; Pred. No. 2.5e-16;
Matches 106; Conservative 86; Mismatches 169; Indels 141; Gaps 16;
Qy 42 EERTESRHFTSKPVLLD-----NILQIRHTPHNGLSDIYLLDDPHEALAAALIES 95
Db 105 DNKOVQKHDLVRMLMDQDGLTENKVDHFDIGNDLYDQVLKD-----IKN 152
Qy 96 AEHSLDLYQYIWRNDISGRLLFNLMYLAERGVRVRLLDNNTNRGLDLDLALDHPNI 155
Db 153 AKEVIHLEYTFALDGLGKILHALEKLGKLEKLVKLYDDVGSK-----NV 199
Qy 156 EVRLNFPVLRKWRALG-----YLTDFP-----RLNRRMNKSFADNRATILGGRNIGD 205
Db 200 KMANFDHF-----KSLGGEVEAFASKPLPLNFRMNRNRKIIIVIDGQLGVYGVGNIGD 254
Qy 206 EYFKVGEDTVFADLDILATGVSVEVSHDFDRYWAHSAHNATRII-----RSGNIGK 258
Db 255 EYLGKLGKLYWRDTHRIQGDALQRLFDLWNS--QAHRPQFEYDVVKYFPKNGPLG- 312
Qy 259 GLQALGYNDTSRHALLRYRETVEQSPLYQKIOTGRIDWQSVQ---TRLISDTPAKGLDR 315
Db 313 -----NSPIQTAASGPASDWHQIEGYTKMMSA----- 341
Qy 316 DRKPPPIAGRLQDALQKPEKSVLYSPYFPTKSGTDALAKLVODGIDVTVLNLSQATD 375
Db 342 -----KKSIVLQSPYFIPDINSYINAIKIAKSGVDVHLMIPC--KPD 381
Qy 376 VAAVHSGYKVRKPLKAGIKLYELQPNHAVPATKDKGLTGSSVTSLSHAKTIFVDGKRIF 435
Db 382 HPLVYWATFNSADLLSSGVKIY-----TYENG-----IHSKMLIDDEIVS 424
Qy 436 IGSFNLDPRGARLNTGEMVWVIESPKIAEQMERTLADTTPVAYRVTLDKHNRQWHPAT 495
Db 425 VGTANMDFRSEFELNFEVNAFVDENLAKDL-----RVAYEHDITKSKQL-----T 469
Qy 496 RKTYPNEPEAKLWKRIAAKILS 517
Db 470 KESYANRPLSVKPKESLAKLVS 491

RESULT 18
US-10-429-094-84
; Sequence 84, Application US/10429094
; Publication No. US20030180821A1
; GENERAL INFORMATION:
; APPLICANT: TOMICH, Che-Shen
; APPLICANT: QUINN, Cheryl
; APPLICANT: ARVIDSON, Staffan
; APPLICANT: HARRIS, Douglas
; APPLICANT: MOTT, John
; TITLE OF INVENTION: ANTIMICROBIAL METHODS AND MATERIALS
; FILE REFERENCE: 00774.US1 CN1
; CURRENT APPLICATION NUMBER: US/10/429,094
; CURRENT FILING DATE: 2003-05-02
; PRIOR APPLICATION NUMBER: US 09/966,521
; PRIOR FILING DATE: 2001-09-28
; NUMBER OF SEQ ID NOS: 145
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 84
; LENGTH: 502
; TYPE: PRT

```

; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Amino acid sequence encoded by S. aureus coding region cloned for
; OTHER INFORMATION: expression in E. coli
US-10-429-094-84

Query Match          10.5%; Score 286.5; DB 4; Length 502;
Best Local Similarity 21.1%; Pred. No. 2.5e-16; Mismatches 169; Indels 141; Gaps 16;
Matches 106; Conservative 86;

Qy 42 EERTESRFTNSKPVLLD-----NILQIRTPHNHGLSDIYLLDDPHEALAAARAAALIES 95
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 96 AEHSLLDQYIWRNDISGRLLFNLMYLAARGVRVRLLLDDNNTRGLDLLLLALDSHPNI 155
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 153 AKEVIHLEYTFALDGLGKRILHALEBKLGQLEVKILYDDVGSK-----NV 199
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 156 EVRLFPVLRKWRALG-----YLTFDP-----RLNRHMNKSFTADNRATILGRNIGD 205
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 200 KMANFDHF-----KSLGGEVEAFFASKLPLNFRNNNNHRRKIIVDGQLYGVGFGNIGD 254
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 206 EYFKVGEDTVFADLDIATGSGVWGEVSHDFRYWASHSAHNATRII-----RSGNIGK 258
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 255 EYLGKLGKGYWRDTHLRIOGDAVDALQLRFILDWNS-QAHRPQFQYDVKYFPKNGPLG- 312
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 259 GLQALGVNDETSRHALLRYETVRSQSPLYQKIQTGRIDMQSVQ---TRLISDTPAKGLDR 315
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 313 -----NSPTQIAASGASPDWHQIEGYCTQVMSA----- 341
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 316 DRRKPPIAGRLQDALKQPEKSVLYLSPYVPTKSGTDALAKLVQDGDIVDTVLTNSLQATD 375
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 342 -----KKSVYLOSPYFIPDINSYINAIKAAGSGVDVHLMIPC--KPD 381
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 376 VAAHSGVYKYRKLKAGIKLYELQPNHAVPATKDKGLTSGSVTSLHAKTFIVDGKRIF 435
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 382 HPLVYVWATFSNASDLLSSGVKIY-----IVENGF-----IHSKWLIDDEIVS 424
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 436 IGSFNLDPRLNTMGVGVIESPKIAEQMERTLADTTPEYAYRVTLDKHNLQWHDHPAT 495
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 425 VGTANMDFRSPFLNFAVFNAYDENLAKDL-----RVAYEHDITKSKQL-----T 469
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 496 RKTYPNEPEAKLWKRIAAKILS 517
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 470 KESYANRPLSVKFESLAKLVS 491
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 19
US-10-335-977-9083
; Sequence 9083, Application US/10335977
; Publication No. US20040052799A1
; GENERAL INFORMATION:
; APPLICANT: DOUGLAS SMITH et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES
; RELATING TO HELICOBACTER PYLORI FOR
; DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 10031
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 28 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: Windows NT 4.0
; SOFTWARE: UNIX
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/335,977
; FILING DATE: 30-Dec-2002
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/993,002
; FILING DATE: 17-DEC-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Mandragouras, Amy E.
; REGISTRATION NUMBER: 36,207
; REFERENCE/DOCKET NUMBER: GTN-018
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)742-4214
; INFORMATION FOR SEQ ID NO: 9083:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 206 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: Helicobacter pylori
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (B) LOCATION 1...206
; SEQUENCE DESCRIPTION: SEQ ID NO: 9083:
US-10-335-977-9083

Query Match          10.1%; Score 274.5; DB 4; Length 206;
Best Local Similarity 31.4%; Pred. No. 7.6e-16;
Matches 64; Conservative 45; Mismatches 84; Indels 11; Gaps 5;

Qy 321 PIAGRLQDALKQPEKSVLYLSPYVPTKSGTDALAKLVQDGDIVDTVLTNSLQATDVAHV 380
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 381 SGYVYKRPKLKAGIKLYELQPNHAVPATKDKGLTSGSVTSLHAKTFIVDGKRIFIGSFN 440
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 66 GAWERYENKLRVMGANYEIRNDFNFRQIKGR---FSTKSLHGKTIIVFDDALTLLGSFN 122
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 441 LDPRSARLNTMGVGVIESPKIAEQMERTLADTTPEYAYRVTLDKHNLQWHDHPATRKTY- 499
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 123 IDPSAVINTESAVLFNFPFSAKVRSLKDKHA-QQSHLVLYRH-RVIWE--ATEGIL 178
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 500 ---ENEPEAKLWKRIAAKILSLLP 520
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 179 IHEKNSPDTSPFLSLIKESWVKLP 202
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 20
US-10-724-972A-4379
; Sequence 4379, Application US/10724972A
; Publication No. US20040147734A1
; GENERAL INFORMATION:
; APPLICANT: Doucette-Stamm, Lynn
; APPLICANT: Bush, David
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: PATH03-16
; CURRENT APPLICATION NUMBER: US/10/724,972A
; CURRENT FILING DATE: 2003-12-01
; PRIOR APPLICATION NUMBER: 09/450,969
; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: 09/134,001
; PRIOR FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 7544
; SEQ ID NO 4379
; LENGTH: 494
; TYPE: PRT
; ORGANISM: S.epidermidis
US-10-724-972A-4379

Query Match          9.5%; Score 259.5; DB 4; Length 494;

```

Best Local Similarity 22.3%; Pred. No. 6.2e-14;
Matches 101; Conservative 70; Mismatches 178; Indels 103; Gaps 11;
Qy 76 IYLLDDPHEALAAARAAALIESAEHSLDQYIWRNDISGRLLFNLMYLAARGVRLD 135
Db 133 IDLFTDGHKLEKLEVDYNAQDYIHLEYTFELDGLGKRLDALETKLEGLVEKLVLD 192
Qy 136 DNNTRGLDLLLLLALDHPNIEVRLNPFVLRKWRALGYLTD--PP-----RLNRRMH 185
Db 193 DVGSK-----KVRLSK---FKHFRALGGEVEAFPPSKVPLINFRMNRNH 234
Qy 186 NKSTADNRATILGRNIGDEYFKVGEDTVPADLDILATGSVGEVSHDPRYWAHSAH 245
Db 235 RKIIIDGQIGYGVGFNGVDYGLGKLGWRDTHTRVQECIDALQLRFLDWNSSOS-- 292
Qy 246 NATRIIRSGNIGKLOALGYNDETSRHALLRYRETVEQSPLYQKIOTGRIDWQSVQRLI 305
Db 293 -----HRPQFKDQYFPKKG--DKGNAALQIA 319
Qy 306 SDTPAKGLDRRRKPPIAGRLQDALQKPEKSVYLVSPYFTKSGTDALAKLVQDGIDVT 365
Db 320 SSGPAFDLHQ-----IEYGYTKMISAKKSIYLSQSPYFIPDQSYINALKWAANSQVEVN 373
Qy 366 VLTNSLOATVAAVHSGYVYKRPGLKAGIKLYELQPNHVPATKDKGLTGSSVTSLSHAK 425
Db 374 LMIPC--KPDHPFYVWATFNAADLDSGVNIYTYQNGF-----IHSK 414
Qy 426 TFIVDGKRIPTGSNLPDRARLNTGEMGVVIESPKIAEQMERTLADTTPYVAVRVTLDKH 485
Db 415 ILMIDDEISSGNMDFRFLNFVNAFYEDIAKLRQ-----AFKEDIEQS 465
Qy 486 NRLQWHDPAKRTKYPNEPEAKLWKRIAAKILS 517
Db 466 KLL-----TKEVYDKRPLSIKFKEGKLAKLS 491

RESULT 21

US-10-627-476-78
; Sequence 78, Application US/10627476
; Patent No. US20040030116A1
; GENERAL INFORMATION:
; APPLICANT: Pompejus, Mark
; APPLICANT: Kroger, Burkhard
; APPLICANT: Schoder, Hartwig
; APPLICANT: Zelder, Oskar
; APPLICANT: Haberhauer, Gregor
; TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING PROTEINS
; TITLE OF INVENTION: INVOLVED IN MEMBRANE SYNTHESIS AND MEMBRANE
; FILE REFERENCE: BGI-125CPCN
; CURRENT APPLICATION NUMBER: US/10/627,476
; CURRENT FILING DATE: 2003-07-25
; PRIOR FILING DATE: 09/602,787
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: USSN 60/141031
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: DE 19931454.3
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19931478.0
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19931563.9
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19932122.1
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932124.8
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932125.6
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932128.0
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932180.9
; PRIOR FILING DATE: 1999-07-09
; Remaining Prior Application data removed - See File Wrapper or PALM.

; NUMBER OF SEQ ID NOS: 678
; SEQ ID NO 78
; LENGTH: 422
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
; US-10-627-476-78

Query Match 9.3%; Score 252; DB 4; Length 422;
Best Local Similarity 23.4%; Pred. No. 2.3e-13;
Matches 108; Conservative 58; Mismatches 179; Indels 116; Gaps 17;

Qy 67 TPNHNGLSDIYLLDDPHEALAAARAAALIESAEHSLDQYIWRNDISGRLLFNLMYLAAR 126
Db 38 TGGNNG-----FYSDYRESLKRWTAIDEAEYIYVEIYIMAWDSYTQPPFAALERAHNR 92
Qy 127 GVRVRLD-----DNNTRGLDLLLLLALDHPNIEVRLNPFVLRKWRAL 171
Db 93 GVKVRLLLFDHVGSKVPGYHRLKKELNRMGFAYLM-----LPLQWR-- 135
Qy 172 GYLTDPPRLNRRMHKSFATADRATILGGRN-IGDEY-----PKVGEDTVPADLDILATG 225
Db 136 ---RRFRPDLNRHKKWLIIDGHTAPMGSQLIAPSLQKKNIKLGRE--WKDLMLVELTG 190
Qy 226 SVVGEVSHDPRYWAHSAHNAIRIRSGNIGKLOALGYNDETSRHALLRYRETVEQSP 285
Db 191 PIVSSMEMIFAGDWYVESN-----EALDIRDHAEAHGYIGNTKQDSATN 234
Qy 286 LYQKIOTGRIDWQSVQTRLISDTPAKGLDRDRKPPIAGRLQDALQKPEKSVYLVSPYFV 345
Db 235 LVQLIPSG-----PGYTTEPNLR-----MFSIVHHAKERLILCSPYFI 273
Qy 346 PTKSGTDALAKLVQDGIDVTVLTNSLOATVAAVHSGYVYKRPGLKAGIKLYELQPNHA 405
Db 274 PDESLLAEVTSACYRGVTVELFVS--EQADQFAIDHAQSSYYQALLEAGVKIYQF-PKPD 330
Qy 406 VPATKDKGLTGSSVTSLSHAKTFIVD-----GKRIF--IGSFNLDPPRSARLNTGEMGVVIES 458
Db 331 V-----LHTKYMIAADPDDTGTNEALGVLGSSNLDIRSFGLNBIISLMIK 375
Qy 459 PKIAEQMERTLADTTPYVAVRVTLDKHNRLOWHDHPATKRTY 499
Db 376 GNLIHEL-NALTDRYRTVSFKLTLDKWNQORSW-----RRRY 410

RESULT 22

US-09-848-726-2
; Sequence 2, Application US/09848726
; Patent No. US20020102667A1
; GENERAL INFORMATION:
; APPLICANT: NAMPOOTHIRI, Madhavan
; TITLE OF INVENTION: No. US20020102667A1e1 Nucleotide Sequences Coding for the cls
; FILE REFERENCE: 032301 WD 1191
; CURRENT APPLICATION NUMBER: US/09/848,726
; CURRENT FILING DATE: 2001-05-04
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 500
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
; US-09-848-726-2

Query Match 9.3%; Score 252; DB 3; Length 500;
Best Local Similarity 23.4%; Pred. No. 3e-13;
Matches 108; Conservative 58; Mismatches 179; Indels 116; Gaps 17;

Qy 67 TPNHNGLSDIYLLDDPHEALAAARAAALIESAEHSLDQYIWRNDISGRLLFNLMYLAAR 126
Db 116 TGGNNG-----FYSDYRESLKRWTAIDEAEYIYVEIYIMAWDSYTQPPFAALERAHNR 170
Qy 127 GVRVRLD-----DNNTRGLDLLLLLALDHPNIEVRLNPFVLRKWRAL 171
Db 171 GVKVRLLLFDHVGSKVPGYHRLKKELNRMGFAYLM-----LPLQWR-- 213

GENERAL INFORMATION:
; APPLICANT: MCMICHAEL, JOHN CALHOUN
; APPLICANT: ZAGURSKY, ROBERT JOHN
; APPLICANT: RUSSELL, DAVID PARRISH
; APPLICANT: FLETCHER, LEAH DIANE
; TITLE OF INVENTION: ALLOIOCOCCUS OTITIDIS OPEN READING FRAMES (ORFS) ENCODING
; FILE REFERENCE: AM100780 L2
; CURRENT APPLICATION NUMBER: US/10/501,282
; CURRENT FILING DATE: 2004-07-09
; PRIOR APPLICATION NUMBER: 60/333,777
; PRIOR FILING DATE: 2001-11-29
; TITLE OF INVENTION: ALLOIOCOCCUS OTITIDIS OPEN READING FRAMES (ORFS) ENCODING
; FILE REFERENCE: AM100780 L2
; CURRENT APPLICATION NUMBER: US/10/501,282
; CURRENT FILING DATE: 2004-07-09
; PRIOR APPLICATION NUMBER: 60/333,777
; PRIOR FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: 60/426,742
; PRIOR FILING DATE: 2002-11-18
; PRIOR APPLICATION NUMBER: PCT/US02/36123
; PRIOR FILING DATE: 2002-11-25
; NUMBER OF SEQ ID NOS: 6653
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2386
; LENGTH: 448
; TYPE: PRT
; ORGANISM: Alloioococcus otitidis
US-10-501-282-2386

Query Match 7.5%; Score 205; DB 5; Length 448;
Best Local Similarity 21.1%; Pred. No. 3.9e-09;
Matches 92; Conservative 64; Mismatches 151; Indels 130; Gaps 17;

QY	80	DDPHEALAAARAALIESAEHSLDLOYYIWRNDISGRLLFNLMLAAERGVRVRLLLDD--NN	138
Db	90	EDLFEALKED---LRQAKYIFLEFYILEKGTMLDEILELFDKQAQGLDIRLLYDDVGN	146
QY	139	TRGLDDLLALDGHNPNIIEVRLFPVLRKWRALGYLTFPRLNRRMHNKSTADNRATIL	198
Db	147	ILRIDENFVSYLENEKIKTSVFNPL---DWR---LTF---QYNYVRDHKIMVVDGKIGYT	197
QY	199	GGRNIGDEYF---KVGE-----DTVFADLDILATGSSVVGVEVSHDF---	235
Db	198	GGMNIGDYNIRVEKAGHWKGGIRLEGQGVWGFTTMF-----LSLWDYLNDDTQDFRDF	252
QY	236	--DRYWASHAHNATRIIRSGNIGKGLQALGY-----NDTSRHALLRYRETVEQSPLYQ	288
Db	253	HPDEY-----KSAENVVTPG-----GYVHPFADDPNRIQI-----SESLYL	289
QY	289	KIOTGRIDMQSVQTRLISDTPAKGLDRDRKPPIAGRLQDALKQPEKSVLVSPYFVPTK	348
Db	290	KL-----IYNAKESIYVKTPYLIFSQ	310
QY	349	SGTDALAKLVQDIDVTVLNSLOATDVAHVHSGYVKYKPKLKAGIKLYELOPNHAVPA	408
Db	311	KLYSALENAALSGVDVRIIVTPGI--PDKKIVFETTQSFYDKLLEVGKVIYVAPGF---	364
QY	409	TKDKGLTGSSTVSLHAKTFIVDGKRIFGSNLDPRSARLNTMGVVIESPKEAQWERT	468
Db	365	-----IHEKVIIDQDFAINGTINFYRSLHHSFECGVLFYNTQSIIDMKND	411
QY	469	LADTTPAYRVRTLDKH	485
Db	412	FDNLFP-ICRQVSLEEN	427

RESULT 30
US-10-501-282-2386
; Sequence 2386, Application US/10501282
; Publication No. US20050203280A1
; GENERAL INFORMATION:
; APPLICANT: MCMICHAEL, JOHN CALHOUN
; APPLICANT: ZAGURSKY, ROBERT JOHN
; APPLICANT: RUSSELL, DAVID PARRISH
; APPLICANT: FLETCHER, LEAH DIANE
; TITLE OF INVENTION: ALLOIOCOCCUS OTITIDIS OPEN READING FRAMES (ORFS) ENCODING
; FILE REFERENCE: AM100780 L2
; CURRENT APPLICATION NUMBER: US/10/501,282
; CURRENT FILING DATE: 2004-07-09
; PRIOR APPLICATION NUMBER: 60/333,777
; PRIOR FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: 60/426,742
; PRIOR FILING DATE: 2002-11-18
; PRIOR APPLICATION NUMBER: PCT/US02/36123
; FILE REFERENCE: AM100780 L2

; CURRENT APPLICATION NUMBER: US/10/501,282
; CURRENT FILING DATE: 2004-07-09
; PRIOR APPLICATION NUMBER: 60/333,777
; PRIOR FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: 60/426,742
; PRIOR FILING DATE: 2002-11-18
; PRIOR APPLICATION NUMBER: PCT/US02/36123
; PRIOR FILING DATE: 2002-11-25
; NUMBER OF SEQ ID NOS: 6653
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2388
; LENGTH: 481
; TYPE: PRT
; ORGANISM: Alloioococcus otitidis
US-10-501-282-2388

Query Match 7.5%; Score 205; DB 5; Length 481;
Best Local Similarity 21.1%; Pred. No. 4.3e-09;
Matches 92; Conservative 64; Mismatches 151; Indels 130; Gaps 17;

QY	80	DDPHEALAAARAALIESAEHSLDLOYYIWRNDISGRLLFNLMLAAERGVRVRLLLDD--NN	138
Db	123	EDLFEALKED---LRQAKYIFLEFYILEKGTMLDEILELFDKQAQGLDIRLLYDDVGN	179
QY	139	TRGLDDLLALDGHNPNIIEVRLFPVLRKWRALGYLTFPRLNRRMHNKSTADNRATIL	198
Db	180	ILRIDENFVSYLENEKIKTSVFNPL---DWR---LTF---QYNYVRDHKIMVVDGKIGYT	230
QY	199	GGRNIGDEYF---KVGE-----DTVFADLDILATGSSVVGVEVSHDF---	235
Db	231	GGMNIGDYNIRVEKAGHWKGGIRLEGQGVWGFTTMF-----LSLWDYLNDDTQDFRDF	285
QY	236	--DRYWASHAHNATRIIRSGNIGKGLQALGY-----NDTSRHALLRYRETVEQSPLYQ	288
Db	286	HPDEY-----KSAENVVTPG-----GYVHPFADDPNRIQI-----SESLYL	322
QY	289	KIOTGRIDMQSVQTRLISDTPAKGLDRDRKPPIAGRLQDALKQPEKSVLVSPYFVPTK	348
Db	323	KL-----IYNAKESIYVKTPYLIFSQ	343
QY	349	SGTDALAKLVQDIDVTVLNSLOATDVAHVHSGYVKYKPKLKAGIKLYELOPNHAVPA	408
Db	344	KLYSALENAALSGVDVRIIVTPGI--PDKKIVFETTQSFYDKLLEVGKVIYVAPGF---	397
QY	409	TKDKGLTGSSTVSLHAKTFIVDGKRIFGSNLDPRSARLNTMGVVIESPKEAQWERT	468
Db	398	-----IHEKVIIDQDFAINGTINFYRSLHHSFECGVLFYNTQSIIDMKND	444
QY	469	LADTTPAYRVRTLDKH	485
Db	445	FDNLFP-ICRQVSLEEN	460

RESULT 31
US-10-501-282-2390
; Sequence 2390, Application US/10501282
; Publication No. US20050203280A1
; GENERAL INFORMATION:
; APPLICANT: MCMICHAEL, JOHN CALHOUN
; APPLICANT: ZAGURSKY, ROBERT JOHN
; APPLICANT: RUSSELL, DAVID PARRISH
; APPLICANT: FLETCHER, LEAH DIANE
; TITLE OF INVENTION: ALLOIOCOCCUS OTITIDIS OPEN READING FRAMES (ORFS) ENCODING
; FILE REFERENCE: AM100780 L2
; CURRENT APPLICATION NUMBER: US/10/501,282
; CURRENT FILING DATE: 2004-07-09
; PRIOR APPLICATION NUMBER: 60/333,777
; PRIOR FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: 60/426,742
; PRIOR FILING DATE: 2002-11-18
; PRIOR APPLICATION NUMBER: PCT/US02/36123
; FILE REFERENCE: AM100780 L2


```
Best Local Similarity 20.4%; Pred. No. 8.9e-07;
Matches 83; Conservative 44; Mismatches 122; Indels 158; Gaps 12;

QY 76 IYLLDDPHEALAAARALIESAEHSLDLQYYIWRNDISGRLLFNLMYLAERGVRLLLD 135
Db 11 IQLENGEQYYPVAFKAIGEAQERIIILETFIWFEDDVKGQJHAALLAAQGVKAEVLDD 70
QY 136 DNNTRGLDLLLLALDSDHPNIEVRLNPFVLRKWRALGYLTDFFPRLNRRMINKSFTADNRA 195
Db 71 GYGSPDLSDEFVNELTAAGVVFRIYDP-----RPLFGWRTN---VFRMRHKIIVIDARI 123
QY 196 TILGGRNIGDEYFKVGEDTVFADLDILATGVSUVGEVSHDFD-----RYWASHSA 244
Db 124 AFIGGLNYSAEHMSYSGPEAKQDYAVRLEGPVIEDIL-QFELENLPQCSAARRWRRH-- 180
QY 245 HNATRIIRSGNIGKGLQALGYNDTSRHALLRYRETVEQSPLYQKIQTGRIDWQSVQTRL 304
Db 181 HKAENRQPE-----AQANGY-----AVRF----- 202
QY 305 ISDTPAKGLDRDRKPPAGRLQDALQKPEKSVYLVSPYFVPTKSGTDLAKLVQDGDIV 364
Db 203 -----GRL----- 205
QY 365 TVLTNSLOATDVAAGHSGYVKYRKPILLKAGIKLYELQPNHAVPATKDKGLTGSSVTSLSHA 424
Db 206 -----SGYY-----LVKGGVQVFEYRR-----PLHG 227
QY 425 KTFIVDGKRIFGSNLDPR SARLNTMGVYVIESPKIAEQMERTLAD 471
Db 228 KVALMDDHWATVGSSNLDPLSLNLEANVHID-----RHFNOTLDR 270

RESULT 34
US-10-450-763-31510
; Sequence 31510, Application US/10450763
; Publication No. US20050196754A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 790CIP3/US
; CURRENT APPLICATION NUMBER: US/10/450,763
; CURRENT FILING DATE: 2003-06-11
; PRIOR APPLICATION NUMBER: PCT/US01/08631
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 31510
; LENGTH: 1097
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (456)..(476)
; OTHER INFORMATION: Phospholipase D active site proteins motifs. domain
; OTHER INFORMATION: identified by eMATRIX, accession number PF00614B, p-value=4.774e-
; OTHER INFORMATION: 14, raw score of 14.45
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (613)..(714)
; OTHER INFORMATION: Sodium:sulfate symporter transmembran domain identified by
; OTHER INFORMATION: Pfam, accession name Na_sulph_syp, E-value=6.1e-46, Pfam score c
; OTHER INFORMATION: 163.9
US-10-450-763-31510

Query Match 6.4%; Score 175; DB 5; Length 1097;
Best Local Similarity 20.4%; Pred. No. 7.1e-06;
Matches 83; Conservative 44; Mismatches 122; Indels 158; Gaps 12;

QY 76 IYLLDDPHEALAAARALIESAEHSLDLQYYIWRNDISGRLLFNLMYLAERGVRLLLD 135
```

```
Db 241 IQLENGEQYYPVAFKAIGEAQERIIILETFIWFEDDVKGQJHAALLAAQGVKAEVLDD 300
QY 136 DNNTRGLDLLLLALDSDHPNIEVRLNPFVLRKWRALGYLTDFFPRLNRRMINKSFTADNRA 195
Db 301 GYGSPDLSDEFVNELTAAGVVFRIYDP-----RPLFGWRTN---VFRMRHKIIVIDARI 353
QY 196 TILGGRNIGDEYFKVGEDTVFADLDILATGVSUVGEVSHDFD-----RYWASHSA 244
Db 354 AFIGGLNYSAEHMSYSGPEAKQDYAVRLEGPVIEDIL-QFELENLPQCSAARRWRRH-- 410
QY 245 HNATRIIRSGNIGKGLQALGYNDTSRHALLRYRETVEQSPLYQKIQTGRIDWQSVQTRL 304
Db 411 HKAENRQPE-----AQANGY-----AVRF----- 432
QY 305 ISDTPAKGLDRDRKPPAGRLQDALQKPEKSVYLVSPYFVPTKSGTDLAKLVQDGDIV 364
Db 433 -----GRL----- 435
QY 365 TVLTNSLOATDVAAGHSGYVKYRKPILLKAGIKLYELQPNHAVPATKDKGLTGSSVTSLSHA 424
Db 436 -----SGYY-----LVKGGVQVFEYRR-----PLHG 457
QY 425 KTFIVDGKRIFGSNLDPR SARLNTMGVYVIESPKIAEQMERTLAD 471
Db 458 KVALMDDHWATVGSSNLDPLSLNLEANVHID-----RHFNOTLDR 500

RESULT 35
US-10-474-792-354
; Sequence 354, Application US/10474792
; Publication No. US20040236072A1
; GENERAL INFORMATION:
; APPLICANT: Olmsted, Stephen
; APPLICANT: Zagursky, Robert
; APPLICANT: Nickbarg, Elliot
; APPLICANT: Winter, Lourie
; TITLE OF INVENTION: SURFACE PROTEINS OF STREPTOCOCCUS PYOGENES
; FILE REFERENCE: AM 100399
; CURRENT APPLICATION NUMBER: US/10/474,792
; CURRENT FILING DATE: 2003-10-14
; NUMBER OF SEQ ID NOS: 674
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 354
; LENGTH: 525
; TYPE: PRT
; ORGANISM: Streptococcus pyogenes
US-10-474-792-354

Query Match 6.2%; Score 168.5; DB 5; Length 525;
Best Local Similarity 20.0%; Pred. No. 8.8e-06;
Matches 89; Conservative 63; Mismatches 131; Indels 163; Gaps 20;

QY 96 AEHSLDQYYI-----WRNDISGRLLFNLMYLAERGVRLLLDNNTRGLDLDL-LLA 148
Db 185 AKKYIFLEFFIIAEGQMW-----GEIL-SILEKKVSEGVVRVLF-----GNELSTLS 233
QY 149 LD-----SHPNIEVRLF---NPFVLRKWRALGYLTDFFPRLNRRMINKSFTADNRAITLGG 200
Db 234 SDYAKELEIQIGIKAKSFLPISPFISTYY-----NYDRHRKIVVIDGEVSTGG 281
QY 201 RNIGDEYF-KVGEDTVFADLDILATGVS-----VGEVSHDFRYWASHSAHN 246
Db 282 INLADYIYNEVERFHWKAGLMEGEATDSFILFLQWMSITEKEKLIIDPYLSDSLK- 340
QY 247 ATRIIRSGNIGKGLQALGYNDTSRHALLRYRETVEQSPLYQKIQTGRIDWQSVQTRLIS 306
Db 341 -----LPSDGY-----VIPY----- 351
QY 307 DTPAKGLDRDRKPPAGRLQDALQKPEKSVYLVSPYFVPTKSGTDLAKLVQDGDIV 364
Db 352 DSP-----LDTDK-----IGKNYVIDILNHAKEYVYIMTPYLIILDSEMEHALRFASERGVDI 403
```



```

RESULT 38
US-10-498-327-79
: Sequence 79, Application US/10498327
: Publication No. US20050106162A1
: GENERAL INFORMATION:
: APPLICANT: Grandi, Giulio
: APPLICANT: Ratti, Giulio
: TITLE OF INVENTION: Immunisation Against Chlamydia Trachomatis
: FILE REFERENCE: 002441.00085
: CURRENT APPLICATION NUMBER: US/10/498,327
: CURRENT FILING DATE: 2004-06-10
: PRIOR APPLICATION NUMBER: PCT/IB02/05761
: PRIOR FILING DATE: 2002-12-12
: NUMBER OF SEQ ID NOS: 262
: SOFTWARE: PatentIn version 3.2
: SEQ ID NO 79
: LENGTH: 474
: TYPE: PRT
: ORGANISM: Chlamydia trachomatis
US-10-498-327-79

Query Match          4.7%; Score 127; DB 5; Length 474;
Best Local Similarity 20.3%; Pred. No. 0.038;
Matches 98; Conservative 89; Mismatches 194; Indels 102; Gaps 23;

Qy 66 HT----PHNNGLSDIYLLDDPHEALAAARAALTESAHSLSLQYYIWRNDISGRLLFNLMYL 122
||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
Db 24 HTTAAIPDGDKKAKVLIIHDNGYMEYHLLAAISSAKYTVELCPCLAGGEILSTVLQRLER 83
||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :

Qy 123 AARGVVRVRLLL-----DDNNTRGLLDLLLLASHNIEVRLFNPFVLRKWRALGYLTD 176
||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
Db 84 MEEVPALVSYLVQPTCIDDDNRKNLKT-----QENYPD-----RFF--YLFSDMPPYCNVF- 134
||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :

Qy 177 FPLNRMRMHNSFTADNRATILGGRTIGDEYFKVGEDTVFADLDIL-ATGSWGEVSHDF 235
||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
Db 135 FPNVT-BSHTKLSIVDGKYIFIGSNLEDLQCSKGD---VDLEVSDFSRAVIGGLRP- 188
||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :

Qy 236 DRYWASHSAHNATRIIRSGNIKGKLGALGYNDETSRHALLRYRTVQSPLYQKIQTGRI 295
||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
Db 189 -----SAMEDQDVTIVSEBYGALLRKEFCAHYAL--WKDFTQKLMNKKLD---- 232
||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :

Qy 296 DWOSVQRLISDTPAKG-----LDR-----DRKPPITAGRLQDALK 331
||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
Db 233 DFRGIDPINLSIEKARSFCAMIETSLCAVSVPLDKMHFIFSGPDESNNTIAEYVRLIN 292
||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :

Qy 332 QPEKSVYLVSPYFVPTKSGTDALAKLVQD-GIDVTVLTNSL--QATDVAAVHS--GYVKY 386
||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
Db 293 QAQHSIRIAQWFFIPVAKIYDLSLMAACWDGRGVEIYLVNTRGTRDSPEITRSYAMGNRINY 352
||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :

Qy 387 -----RKPLL-----KAGIKLYELQPNHAVPATKGLTGSSVTSLSHAKTFIVDG 431
||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
Db 353 PPIITFGSRPLLWERFLYSPGRASMKFYVSEFYA-----NTQLHKKCMVLVDD 399
||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :

Qy 432 KRIFIGSNLDPRSARLNTMGVVIESPKEIAQWERTLADTTPPEYAYRVLTDKHNRLQWH 491
||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
Db 400 HILVIGSYNFGKSNDCDYECIVVIDS-KEAVSKAQVVFKEKDLRLSKSVTHD--DIINWY 456
||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :

Qy 492 -DP 493
|||
Db 457 FDP 459
|||

```

RESULT 39
US-10-408-765A-2203
; Sequence 2203, Application US/10408765A
; Publication No. US20040101874A1
; GENERAL INFORMATION:
; APPLICANT: Ghosh, Soumitra S.
; APPLICANT: Fahy, Eoin D.
; APPLICANT: Zhang, Bing
; APPLICANT: Gibson, Bradford W.
; APPLICANT: Taylor, Steven W.

```

; APPLICANT: Glenn, Gary M.
; APPLICANT: Warnock, Dale E.
; TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION
; FILE REFERENCE: 660088.465
; CURRENT APPLICATION NUMBER: US/10/408,765A
; CURRENT FILING DATE: 2003-04-04
; NUMBER OF SEQ ID NOS: 3077
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2203
; LENGTH: 548
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-408-765A-2203

Query Match          4.6%; Score 124.5; DB 4; Length 548
Best Local Similarity 18.0%; Pred. No. 0.079;
Matches 96; Conservative 76; Mismatches 145; Indels 2
Qy 124 AEGGVRRVLLDDNNTRGLDOLLALDS-----HPNIEVRLFNFPFQ
Db 6 AEEGVRSILLFKE-----VELALGINSGYSKRALMLLPNIKV-MRHPDQ
Qy 174 LTDEPLNRNRHMKSFADNRATILGG-----RNIGD-----
Db 55 -----AHHEKLLVVDQVAFGLGLDLAYGRWDDLHYRLTLDLGSSSE
Qy 206 -----EYFKVGED-----TVFLT-----
Db 106 PDSPATPDLSHQFVWLGKDYSLITKDWQLDRPPEDFDIRETTPRMPWPM
Qy 227 VGEVSHDFORYWASHSAHNATRIIRSGNICKGLQALGYNDETSRHALLR
Db 166 PARDLARHFQRM-----NFTKTTKA-----KYKTPPYLLLPKP
Qy 287 YKIQIGRIDWQSVQTHLISDTPAKGLDRDRRKPIIAGRLQDA
Db 210 --TLPGGQC--TTVQV-----LRSVDRWS-----AGTLENSILNAYLHT
Qy 339 LVSPYFVPTKSGTDALAKLYQDGDID-----VTVLT-----
Db 254 IENQFFISCSGDRITLVKNKGDEIVDRILKAHQKGCYRVVYVLLPPLPGFGE
Qy 371 LQATDVAHVHSGY-----VKYRKPLLKAGIKLYELQ
Db 314 IQ-----AILHFTYRTLGRGEVSLHLRLKAANGTAWRDYISICGLRTHGELG
Qy 412 KGUTGSVTSLHAKTFIVDGKRFIGSFNLDPRS--ARLNTMGVWIE--
Db 365 -----SELIYHKSVLIAADRTVIIGSANINDRSLGLGRDSELAVLIEDTFT
Qy 463 EQMERTLADTTPVAYRVTLDKNR--LQWHDPAKRTYTPNEPEAKLWKR
Db 420 EYQAGRFAULSRKHCFVILGANTRPDLDRDPICDDFF-----QLWQDQD

RESULT 40
US-10-424-599-215690
; Sequence 215690, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 215690
; LENGTH: 553

```

```
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_36797C.1.pap
US-10-424-599-215690

Query Match      4.5%; Score 122.5; DB 4; Length 553;
Best Local Similarity 18.2%; Pred. No. 0.12;
Matches 92; Conservative 79; Mismatches 167; Indels 167; Gaps 21;

QY 116 LFNLMYLAERGVRRLLDDNNT-----RGL----DDLILALDSDHPNIEVRLNPFV 164
Db 26 LGELLKKAEGVAVRVWIDDETSLPFIKKNGVMKTHDEDAFAFKTKV-----I 77

QY 165 LKWRALGYLTDPPRLNRRMHNKSFADNRA-----TILGRNIGDEYFKVGEDT 214
Db 78 CRKCPRLHH--EFPTLFAH-HQKTTITVDTRAPNSVRAREIMSFLLGGVLDLDCGRYDTEKHS 134

QY 215 VFA-----DLILATGSVVGEYSHDFDRYWASHS 243
Db 135 LFQTLTEESHYHDFYQTNAGASLNKGGPREPWHDAHASVTGDAAMDVLTNPEQRWTKQC 194

QY 244 AHNATRIIRSGNIGKGLQALGYNDTSRHALLRYR--ETVEOSPLYQKIQTGRIDMQSVQ 301
Db 195 --DASLLVPANTLENLIPTCSSPPKERNKWQVYRSIDHVSASQLFRKLTVER-----SIH 248

QY 302 TRLISDTPAKGLDRDRKPPPIAGRLQDALKQPEKSVYLVSPYF----- 344
Db 249 EAYV-----EAIIRRADRFVYIENQYFIGGCHLWQKDRHSGCRN 286

QY 345 -VPTKSGTDLAKL-----VODGIDVT--VLINSLQATDVA 377
Db 287 LIPVEIALKVSISKIKARERFAVYIIPMWPEGVPESEPVQDILHWTRTETMINMYKLI 346

QY 378 AVHSGYVKYRKPLL-----KAGIKLYELQPNHAVPAT-----KDKGLTGSSVTS 422
Db 347 IIESEPGHPRDYLNFFCLANREKKGSEY-LSPHSPHPEIQYWNAQKNRFP-----VVV 401

QY 423 HAKTFIVDGKRIFIGSFNLDPRSARLNTMGVVIESPKEIAEQMERTLADTTPEYAYRVT 482
Db 402 HSKIMIVDDIYILIGSANVNQSRMDGQDRTETAIAGAYSQDGADHHIISRGDIHAYRMS 461

QY 483 DKHNLQW--HDPATRKTYPNEPEA 505
Db 462 -----WYEHTGITEELF-LEPES 478
```

Search completed: May 2, 2006, 06:28:22
Job time : 170 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: May 2, 2006, 06:25:45 ; Search time 28 Seconds
(without alignments)
852.710 Million cell updates/sec

Title: US-10-665-990A-14
Perfect score: 2720
Sequence: 1 MHTDPKIQAMPSETISPMKT.....KLMKRIAATILSLPIRGLL 525

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 232119 seqs, 45477862 residues

Total number of hits satisfying chosen parameters: 232119

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA New:
1: /SIDSS5/ptodata/2/pubpaa/US08_NEW_PUB.pep1.*
2: /SIDSS5/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
3: /SIDSS5/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
4: /SIDSS5/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
5: /SIDSS5/ptodata/2/pubpaa/RCT_NEW_PUB.pep.*
6: /SIDSS5/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
7: /SIDSS5/ptodata/2/pubpaa/US09_NEW_PUB.pep1.*
8: /SIDSS5/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
9: /SIDSS5/ptodata/2/pubpaa/US10_NEW_PUB.pep1.*
10: /SIDSS5/ptodata/2/pubpaa/US11_NEW_PUB.pep.*
11: /SIDSS5/ptodata/2/pubpaa/US11_NEW_PUB.pep1.*
12: /SIDSS5/ptodata/2/pubpaa/US60_NEW_PUB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2712	99.7	525	9	US-10-467-657-5462
2	294.5	10.8	504	11	US-11-045-004-2518
3	287	10.6	482	11	US-11-045-004-2698
4	241.5	8.9	503	11	US-11-079-463-7133
5	232	8.5	442	9	US-10-793-626-1000
6	229.5	8.4	429	11	US-11-079-463-7814
7	227.5	8.4	485	11	US-11-074-176-10
8	226.5	8.3	396	9	US-10-510-386-238
9	226.5	8.3	399	9	US-10-510-386-30
10	220	8.1	479	11	US-11-098-686-10838
11	108.5	4.0	1263	9	US-10-485-517-127
12	105.5	3.9	1234	9	US-10-467-657-4224
13	104.5	3.8	489	9	US-10-242-586-42
14	104.5	3.8	489	9	US-10-242-902-42
15	104.5	3.8	489	9	US-10-243-116-42
16	104.5	3.8	489	9	US-10-243-136-42
17	104.5	3.8	489	9	US-10-243-189-42
18	104.5	3.8	489	9	US-10-243-215-42
19	104.5	3.8	489	9	US-10-243-236-42
20	104.5	3.8	489	9	US-10-243-298-42
21	104.5	3.8	489	9	US-10-243-304-42

22	104.5	3.8	489	9	US-10-243-338-42	Sequence 42, Appl
23	104.5	3.8	489	9	US-10-243-345-42	Sequence 42, Appl
24	104.5	3.8	489	9	US-10-243-357-42	Sequence 42, Appl
25	104.5	3.8	489	9	US-10-245-083-42	Sequence 42, Appl
26	104.5	3.8	489	9	US-10-247-015-42	Sequence 42, Appl
27	104.5	3.8	489	11	US-11-157-996-2	Sequence 15, Appl
28	104.5	3.8	506	11	US-11-226-701-15	Sequence 15, Appl
29	104	3.8	328	11	US-11-087-099-7553	Sequence 7553, Ap
30	103	3.8	495	11	US-11-074-176-266	Sequence 266, App
31	101.5	3.7	1072	11	US-11-096-568A-27848	Sequence 27848, A
32	101.5	3.7	1181	11	US-11-096-568A-27847	Sequence 27847, A
33	101.5	3.7	1189	11	US-11-096-568A-27846	Sequence 27846, A
34	101	3.7	833	7	US-09-941-095-85	Sequence 85, Appl
35	101	3.7	833	11	US-11-198-746-85	Sequence 85, Appl
36	101	3.7	833	11	US-11-198-794-85	Sequence 85, Appl
37	101	3.7	833	11	US-11-198-657-85	Sequence 85, Appl
38	101	3.7	1057	11	US-11-096-568A-30809	Sequence 30809, A
39	101	3.7	1166	11	US-11-096-568A-30808	Sequence 30808, A
40	101	3.7	1174	11	US-11-096-568A-30807	Sequence 30807, A
41	101	3.7	1423	11	US-11-188-298-17435	Sequence 17435, A
42	100	3.7	331	11	US-11-045-004-674	Sequence 674, App
43	100	3.7	423	11	US-11-087-099-9468	Sequence 9468, App
44	100	3.7	1159	9	US-10-055-877-139	Sequence 139, App
45	100	3.7	1574	11	US-11-212-443-179	Sequence 179, App

ALIGNMENTS

RESULT 1
US-10-467-657-5462
; Sequence 5462, Application US/10467657
; Publication No. US20050260581A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON SpA
; APPLICANT: FONTANA Maria Rita
; APPLICANT: PIZZA Mariagrazia
; APPLICANT: MASIGNANI Vega
; APPLICANT: MONACI Elisabetta
; TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/467,657
; CURRENT FILING DATE: 2003-08-11
; PRIOR APPLICATION NUMBER: GB-0103424.8
; PRIOR FILING DATE: 2001-02-12
; NUMBER OF SEQ ID NOS: 9218
; SOFTWARE: SeqWin99, version 1.04
; SEQ ID NO 5462
; LENGTH: 525
; TYPE: PRT
; ORGANISM: Neisseria gonorrhoeae
US-10-467-657-5462

Query Match	99.7%	Score 2712;	DB 9;	Length 525;
Best Local Similarity	99.6%	Pred. No. 5.7e-215;		
Matches 523;	Conservative 1;	Mismatches 1;	Indels 0;	Gaps 0;
Qy	1	MHTDPKIQAMPSTISPMKTRSLISLCLLCCSSWLPLEERTESRHNTPKPVLLDN	60	
Db	1	MHTDPKIQAMPSTISPMKTRSLISLCLLCCSSWLPLEERTESRHNTPKPVLLDN	60	
Qy	61	ILQIRHTPHNGLSDIYLLDDPHEALAAARALTESAHSLDLOYYIWRNDISGRLLFNLM	120	
Db	61	ILQIRHTPHNGLSDIYLLDDPHEAPARAALIESAHSLDLOYYIWRNDISGRLLFNLM	120	
Qy	121	YLAERGVRVRLLDNNTRGLDLDLALDSDHNIEVRLFNFPVLRKRWALGYLTFPRL	180	
Db	121	YLAERGVRVRLLDNNTRGLDLDLALDSDHNIEVRLFNFPVLRKRWALGYLTFPRL	180	
Qy	181	NRMHNKSTADNRATILGRNIGDEYFKVGEDTVFADLDILATGSGVGVSHDFORYWA	240	
Db	181	NRMHNKSTADNRATILGRNIGDEYFKVGEDTVFADLDILATGSGVGVSHDFORYWA	240	


```
Qy 241 SHSAHNATRIIRSGNIGKGLQALGYNDTSHALLRYBETVEQSPLYKIQIOTGRIDWQSV 300
Db 241 SHSAHNATRIIRSGNIGKGLQALGYNDTSHALLRYBETVEQSPLYKIQIOTGRIDWQSV 300
Qy 301 QTRLSIDTPAKGLDRDRKPIAGRLQDALKOPEKSVLVSPYFVPTKSGTDLAKLVQD 360
Db 301 QTRLSIDSPAKGLDRDRKPIAGRLQDALKOPEKSVLVSPYFVPTKSGTDLAKLVQD 360
Qy 361 GIDVTVLNSLQATDVAHVSGYVKYKPLKAGIKLYELQPNHAVPATKDKGLTGSSVT 420
Db 361 GIDVTVLNSLQATDVAHVSGYVKYKPLKAGIKLYELQPNHAVPATKDKGLTGSSVT 420
Qy 421 SLHAKTFIVDGKRIFIGSFNLDPRSARLNTGVVIESPKIAEQMERTLADTTPEYAYRV 480
Db 421 SLHAKTFIVDGKRIFIGSFNLDPRSARLNTGVVIESPKIAEQMERTLADTTPEYAYRV 480
Qy 481 TLDKHNRLQWHDPATRKTPYNEPEAKLWKRIAANKILSLPIEGLL 525
Db 481 TLDKHNRLQWHDPATRKTPYNEPEAKLWKRIAANKILSLPIEGLL 525
```

RESULT 2

```
US-11-045-004-2518
; Sequence 2518, Application US/11045004
; Publication No. US20060078901A1
; GENERAL INFORMATION:
; APPLICANT: BUCHRIESER, CARMEN
; APPLICANT: FRANGEUL, LIONEL
; APPLICANT: COUVE, ELISABETH
; APPLICANT: RUSNIOK, CHRISTOPHE
; APPLICANT: FSIHI, HAFIDA
; APPLICANT: DEHOUX, PIERRE
; APPLICANT: DUSSURGET, OLIVIER
; APPLICANT: CHETOUANI, FARID
; APPLICANT: NEDJARI, HAFED
; APPLICANT: GLASER, PHILIPPE
; APPLICANT: KUNST, FRANCK
; APPLICANT: COSSART, PASCALE
; APPLICANT: DANIELS, JUSTIN
; APPLICANT: GOEBEL, WERNER
; APPLICANT: KREFT, JURGEN
; APPLICANT: KUHN, MICHAEL
; APPLICANT: NG, EVA
; APPLICANT: VAZQUEZ-BOLAND, ANTONIO
; APPLICANT: DOMINGUEZ-BERNAL, GUSTAVO
; APPLICANT: GARRIDO-GARCIA, PATRICIA
; APPLICANT: TIERREZ-MARTINEZ, ALBERTO
; APPLICANT: AMEND, ALEXANDRA
; APPLICANT: CHAKRABORTY, TRINAD
; APPLICANT: DOMANN, EUGEN
; APPLICANT: HAIN, THORSTEN
; APPLICANT: BERCHE, PATRICK
; APPLICANT: CHARBIT, ALAIN
; APPLICANT: DURANT, LIONEL
; APPLICANT: PEREZ-DIAZ, JOSE-CLAUDIO
; APPLICANT: BAQUERO, FERNANDO
; APPLICANT: GARCIA DEL PORTILLO, FRANCISCO
; APPLICANT: GOMEZ-LOPEZ, NURIA
; APPLICANT: MADUENIO, ENCARNIA
; APPLICANT: PABLOS, BETRIZ DE
; APPLICANT: WEHLAND, JURGEN
; APPLICANT: KARST, UWE
; APPLICANT: ENTIAN, KARL-DIETER
; APPLICANT: HAUF, JORG
; APPLICANT: ROSE, MATTHIAS
; APPLICANT: VOSS, HAMUT
; TITLE OF INVENTION: LISTERIA MONOCYTOGENES GENOME, POLYPEPTIDES AND USES
; FILE REFERENCE: 05394.0018-02
; CURRENT APPLICATION NUMBER: US/11/045,004
; CURRENT FILING DATE: 2005-01-28
; PRIOR APPLICATION NUMBER: 10/637,657
; PRIOR FILING DATE: 2003-08-11
; PRIOR APPLICATION NUMBER: 10/257,023
```

```
; PRIOR FILING DATE: 2002-10-08
; PRIOR APPLICATION NUMBER: PCT/FR01/01118
; PRIOR FILING DATE: 2001-04-11
; PRIOR APPLICATION NUMBER: FR 00/04,629
; NUMBER OF SEQ ID NOS: 2854
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 2518
; LENGTH: 504
; TYPE: PRT
; ORGANISM: Listeria monocytogenes
US-11-045-004-2518

Query Match 10.8%; Score 294.5; DB 11; Length 504;
Best Local Similarity 23.5%; Pred. No. 4.1e-16;
Matches 119; Conservative 72; Mismatches 207; Indels 109; Gaps 17;

Qy 19 KTRSLISLCLLCCSSWLPLEERTESRHFNTSKPVLNDNLQIRHTPHNGLSDIYL 78
Db 96 KTK-LINAIHAI PNNTNEKLPRLSKRI--AHLTSIEPI-----KG-NKIEI 137
Qy 79 LDDPHEALAAARAAALIESAEHSLDLQYYIWRNDISGRLLFNLMYLAERGVVRLLDDNN 138
Db 138 LTNGEETFPVLLDALRKAENHIHQYVIFKTDALSTEIRDILVEKAKSGVEVRFMDGLG 197
Qy 139 TRGLDDLLALLDHPNIEVRLFNPFVLKWPALCYLTDFFPLNRMRMKNKSTADNRATIL 198
Db 198 SSKLGKAPLAPLKEAGVSIHAFDP-IASPMIV-----RTANLRNHRKIVIDGQIGFT 249
Qy 199 GGRNIGDEY-----FKVGEDTVPADLDILATSGVWGEVSHDFRYWA--SHSAHNATRI 250
Db 250 GGLNIGEYRSNTPDFRVRWDT-----HIKITGAIVELQESFLNDWVYMENQAGAADGF 304
Qy 251 IRSGNIGKGLQALGYNDTSHALLRYRETVEQSPLYKIQIOTGRIDWQSVQTRLSIDTPA 310
Db 305 ISESGSKQYFSPVDMGDWA-----QVIYGGPYD----KEKWVRDS-- 341
Qy 311 KGLDRDRKPIAGRLQDALKOPEKSVLVSPYFVPTKSGTDLAKLVQDGDIVTLVNTS 370
Db 342 -----MLDLSAKESVMVSPYFVPTDEESLAVIRRVAMSGVDVRVITPG 386
Qy 371 LQATDVAHVSGYVKYKPLKAGIKLYELQPNHAVPATKDKGLTGSSVTSIHAFTFIVD 430
Db 387 --KDRGISFHGSNAYVKTMEAGAKMYAYADDSFV-----HAKAMLVLD 428
Qy 431 GKRIFIGSFNLDPRSARLNTGVVIESPKIAEQMERTLADTTPEYAYRVTLDKHNRLQW 490
Db 429 GTRAAIGTANPDFVRSFRLNHELMVFLYDE--SEAMHHLKRDFFKDF-----EDSRL-- 477
Qy 491 HDPATRKTPYNEPEAKLWKRIAANKILS 517
Db 478 ---FTMKDMENKPLLTTRIKEVLSLLS 501
```

RESULT 3

```
US-11-045-004-2698
; Sequence 2698, Application US/11045004
; Publication No. US20060078901A1
; GENERAL INFORMATION:
; APPLICANT: BUCHRIESER, CARMEN
; APPLICANT: FRANGEUL, LIONEL
; APPLICANT: COUVE, ELISABETH
; APPLICANT: RUSNIOK, CHRISTOPHE
; APPLICANT: FSIHI, HAFIDA
; APPLICANT: DEHOUX, PIERRE
; APPLICANT: DUSSURGET, OLIVIER
; APPLICANT: CHETOUANI, FARID
; APPLICANT: NEDJARI, HAFED
; APPLICANT: GLASER, PHILIPPE
; APPLICANT: KUNST, FRANCK
; APPLICANT: COSSART, PASCALE
; APPLICANT: DANIELS, JUSTIN
; APPLICANT: GOEBEL, WERNER
```


Db 157 RVQLLIKKAQGVKRVVVYDAWGSNGATKAWFDQLRKAGGK-----VLPFITSRNM 208
Qy 170 ALGYLTDFPRLNRMRHNSFTADNRATILGRNIGDEYF----KVGEDTVFADLDILATG 225
Db 209 ITRY-----RINYHLHRKIVIDGKISWTGFGNIGDQVLGRKKFGH---WRDSQVRIVG 260
Qy 226 SVGEVSHDFRYWASHAHNATRIIRSGNIGKGLQALGYNDTSRHALLRYRETVEQSP 285
Db 261 SASLLQERFVMDW-NASINNDDEIR-----FN-----ST 290
Qy 286 LYOKIQTGRIDWQSVOTRLISDTAKGLDRDRKPPPIAGRLQDALKOPEKSVYLVSFV 345
Db 291 LFPDLDEKNIHQDDVATILISDGP-----DRYSYMGNGMRLMLARNLWVOTPYLI 344
Qy 346 PTKSGTDALAKLVODGIDVTLTNSLOATDVAASHGVYKVRKPLKAGIKLYELQPNHA 405
Db 345 PDDAVFATWQTIAMSGVDVRIMPC--KPDHPFIYRATQWYANELTRFGVKIYIYEDGF- 401
Qy 406 VPATKDKGLTGSSVTSLSHAKTIFVDGKRIFGSNLDRPSARLNTMGVIE-----SPKI 461
Db 402 -----LHAKTTIIDNFSVSGSMNQDYSYSLNFEDNAIFYDKNFKNKM 445
Qy 462 AEQMERTLADR---TPE 475
Db 446 AEAFEEDMKSHLLTPE 462

RESULT 8
US-10-510-386-238
; Sequence 238, Application US/10510386
; Publication No. US20050244922A1
; GENERAL INFORMATION:
; APPLICANT: Andersen, Jens Tonne
; APPLICANT: Clausen, Ib Groth
; APPLICANT: Jorgensen, Steen Troels
; APPLICANT: Olsen, Peter Bjarke
; APPLICANT: Rasmussen, Michael Dolberg
; TITLE OF INVENTION: Improved Bacillus Host Cell
; FILE REFERENCE: 10294.204-US
; CURRENT APPLICATION NUMBER: US/10/510,386
; CURRENT FILING DATE: 2004-10-04
; NUMBER OF SEQ ID NOS: 248
; SOFTWARE: Patentin version 3.3
; SEQ ID NO 238
; LENGTH: 396
; TYPE: PRT
; ORGANISM: Bacillus licheniformis
US-10-510-386-238

Query Match 8.3%; Score 226.5; DB 9; Length 396;
Best Local Similarity 22.4%; Pred. No. 1.1e-10;
Matches 97; Conservative 58; Mismatches 177; Indels 101; Gaps 11;
Qy 74 SDIYLLDDPHEALAAARALIESAHSLSLDLOYYIWRNDISGRLLFNMLAAERGVRVRL 133
Db 40 SDIELIHNGEDLCERLLDDIRQAESSVHVMFYIKNDDISLEFLKVLKDKAKSGVCVRL 99
Qy 134 LDDNTRGLDLDLALDLSHNIERLFPVLRKWRALGYLTDFPRLNRMRHNSFTADN 193
Db 100 IDRIGAMKVKKTLGSLKQSGVHVFFAN-----KPGPFYFYRLNARNHRKIAVIDG 151
Qy 194 RATILGRNIGDEYF-KVGEDTVFADLDILATGSVVGVEVSH-----DFDRYWASHAHNAT 248
Db 152 KIGVGGFNIKEYLKKAEGFPGWKDYHLRMTGEGVADLQHI FIDPK----- 199
Qy 249 RIIRSGNIGKGLQALGYNDTSRHALLRYRETVEQSP---LYOKIQTGRIDWQSVOTRLI 305
Db 200 -----REAPQAKPANSVFPFLOQGA VTHTHATKGF 230
Qy 306 SDTPAKGLDRDRKPPPIAGRLQDALKOPEKSVYLVSFVPTKSGTDALAKLVQDGIDVT 365
Db 231 S-----LEEKYSFIEQAKERIMICTPYIPSPALQOAVLSARERGIVS 275

Qy 366 VLTNLSLOATDVAASHGVYKVRKPLKAGIKLYELQPNHAVPATKDKGLTGSSVTSLSHAK 425
Db 276 VLVP--MKPDHPLVKEAAYTHFPALLKAGCYIYRY-----RGF-----YHAK 316
Qy 426 TFIVDGKRIFIGSNLDRPSARLNTMGVIESPK-----IAQMERTLADTTPPEYA 477
Db 317 ALIVDDRHRVMIGTSNPNRSLFLNDEVNVVHDKDWTQKQFPDVFVKESIEHAELLTKERYA 376
Qy 478 YRVTLDKHNRLOW 490
Db 377 KRPVMQR--PVEW 387

RESULT 9
US-10-510-386-30
; Sequence 30, Application US/10510386
; Publication No. US20050244922A1
; GENERAL INFORMATION:
; APPLICANT: Andersen, Jens Tonne
; APPLICANT: Clausen, Ib Groth
; APPLICANT: Jorgensen, Steen Troels
; APPLICANT: Olsen, Peter Bjarke
; APPLICANT: Rasmussen, Michael Dolberg
; TITLE OF INVENTION: Improved Bacillus Host Cell
; FILE REFERENCE: 10294.204-US
; CURRENT APPLICATION NUMBER: US/10/510,386
; CURRENT FILING DATE: 2004-10-04
; NUMBER OF SEQ ID NOS: 248
; SOFTWARE: Patentin version 3.3
; SEQ ID NO 30
; LENGTH: 399
; TYPE: PRT
; ORGANISM: Bacillus licheniformis
US-10-510-386-30

Query Match 8.3%; Score 226.5; DB 9; Length 399;
Best Local Similarity 22.4%; Pred. No. 1.1e-10;
Matches 97; Conservative 58; Mismatches 177; Indels 101; Gaps 11;
Qy 74 SDIYLLDDPHEALAAARALIESAHSLSLDLOYYIWRNDISGRLLFNMLAAERGVRVRL 133
Db 43 SDIELIHNGEDLCERLLDDIRQAESSVHVMFYIKNDDISLEFLKVLKDKAKSGVCVRL 102
Qy 134 LDDNTRGLDLDLALDLSHNIERLFPVLRKWRALGYLTDFPRLNRMRHNSFTADN 193
Db 103 IDRIGAMKVKKTLGSLKQSGVHVFFAN-----KPGPFYFYRLNARNHRKIAVIDG 154
Qy 194 RATILGRNIGDEYF-KVGEDTVFADLDILATGSVVGVEVSH-----DFDRYWASHAHNAT 248
Db 155 KIGVGGFNIKEYLKKAEGFPGWKDYHLRMTGEGVADLQHI FIDPK----- 202
Qy 249 RIIRSGNIGKGLQALGYNDTSRHALLRYRETVEQSP---LYOKIQTGRIDWQSVOTRLI 305
Db 203 -----REAPQAKPANSVFPFLOQGA VTHTHATKGF 233
Qy 306 SDTPAKGLDRDRKPPPIAGRLQDALKOPEKSVYLVSFVPTKSGTDALAKLVQDGIDVT 365
Db 234 S-----LEEKYSFIEQAKERIMICTPYIPSPALQOAVLSARERGIVS 278
Qy 366 VLTNLSLOATDVAASHGVYKVRKPLKAGIKLYELQPNHAVPATKDKGLTGSSVTSLSHAK 425
Db 279 VLVP--MKPDHPLVKEAAYTHFPALLKAGCYIYRY-----RGF-----YHAK 319
Qy 426 TFIVDGKRIFIGSNLDRPSARLNTMGVIESPK-----IAQMERTLADTTPPEYA 477
Db 320 ALIVDDRHRVMIGTSNPNRSLFLNDEVNVVHDKDWTQKQFPDVFVKESIEHAELLTKERYA 379
Qy 478 YRVTLDKHNRLOW 490
Db 380 KRPVMQR--PVEW 390

RESULT 10


```
; ORGANISM: Neisseria gonorrhoeae
US-10-467-657-4224

Query Match
Best Local Similarity 3.9%; Score 105.5; DB 9; Length 1234;
Matches 127; Conservative 63; Mismatches 190; Indels 235; Gaps 32;

Qy 39 PPLEERTESRHNTSK---PVLLDNIQLIRHTPHNGLSDIYLLDDPHEALAAALIES 95
Db 4 PIPKPREKSWFNLSQSLPLALARYL-----PHKRL--KAVLTQD 42
Qy 96 AEHSLDLO--YYIWRNDISGRLLFNLMYLAERG-----VRVRLLLDDNNTRGLDLLL 147
Db 43 AEQALRLQTAWRFRPHDHTAVFLPDWETLPHYERFSPHQDLVSRLSALWQIKSGAADVLF 102
Qy 148 ----ALDSDHNIEVRLFNPPVLAK--WRALGYLTDFPLNRM-----HNKSFTADNR 194
Db 103 VPVATAMQKLPV-----PFLAGRTFWLKTGTQTLDIRUKTLDVLDAGYNHSHVVAAGE 156
Qy 195 ATILGGRNIGDEYFKVGEDTVFA-----DLDLATGVSUVGEV-----SHDFD 236
Db 157 FAVREGG--IVD-LFPMGSETPYRIDLPDDEIDSIKTFDTQRTISPVSEIRLLPAHEFP 213
Qy 237 RYASHAHNAIRIIRSGNIKGKGLQALGYNDGTSRHALLRYRTVEOSP-----LYQKIQT 292
Db 214 -----TDSEAQKIFRS-----RPREVDGNPNDAAVYKAVSN 245
Qy 293 GRID-----WQSVQTRLISDTPAK 311
Db 246 GHFGAGVEYIPLPFENELETLFYIGEDALFVSLGVDHAEANRFWNVDKRSY---AMAQ 302
Qy 312 GLDRDRRKPI-----AGRLQD-ALKOPEKSVYLVLS-PYFVPTKSGTDALAKL 357
Db 303 G--DETYPLLPQHLYSADVFAKLNKYGQVLPDVSGRKASHLPDLAVNRQSDPDLOAL 359
Qy 358 --VODGIDVTVL-----TNSLOATDVAAVHSGYVYKVRKPL-----LK 392
Db 360 KDFQTAFDGRILLCAESLGRRETMLGFLQOGLKAKFVSD-WOGFLSAHEPLMTVAPLA 418
Qy 393 AGIKLYELOPN---HAYPATKDKLGTGSSVTSLSHAKTFIVDGKEIFIGSFNLQPRSRALN 449
Db 419 YGFKLGLGSSQQOQTVFASGEKGAVTDQTEFSASA-----TNPLPSPL- 463
Qy 450 TEMGVVIESPKIAEQMERTLAD-----TTPYAYRVTLDKHNLQWHPATKRTKYPNE 502
Db 464 -----PQEREQSAAVSDGLKAAVSTESSLYLVASDLHCQ-----TRQQSAPS 507
Qy 503 PEAKLM-----KRIAAK 514
Db 508 PVGEGWGEKAVAAQ 522
```

```
RESULT 13
US-10-242-586-42
; Sequence 42, Application US/10242586
; Publication No. US20060073548A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin
; APPLICANT: Eaton, Dan
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Goddard, Audrey
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin
; APPLICANT: Smith, Victoria
; APPLICANT: Stephan, Jean-Phillippe
; APPLICANT: Watanabe, Colin
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; APPLICANT: Fong, Sherman
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3630R1C21
; CURRENT FILING DATE: 2002-09-13
; PRIOR APPLICATION NUMBER: US/10/242,586
```

```
; CURRENT FILING DATE: 2002-09-11
; PRIOR APPLICATION NUMBER: 10/197942
; PRIOR FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/059114
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/063046
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/065027
; PRIOR FILING DATE: 1997-11-10
; PRIOR APPLICATION NUMBER: 60/079689
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/086478
; PRIOR FILING DATE: 1998-05-22
; PRIOR APPLICATION NUMBER: 60/087607
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/089801
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/090557
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090689
; PRIOR FILING DATE: 1998-06-25
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 116
; SEQ ID NO 42
; LENGTH: 489
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-242-586-42

Query Match
Best Local Similarity 3.8%; Score 104.5; DB 9; Length 489;
Matches 53; Conservative 33; Mismatches 71; Indels 55; Gaps 11;

Qy 284 SPYIQKIQT-GRIDWQSVQTRLISDTPAKGLDRDRKP-----PIAGRLQDALKOPEKSVYL 339
Db 60 SPAWEPLEAEARQORDSCQVLVLESIP-QDLPSAAGSPSAQPLGQAWLQLLDTAQESVHV 118
Qy 340 VSPFYFVPT-----KSGTDALAKLVQ---DGIDVTVLNS---LQATD--VAAV 379
Db 119 ASYVNSLTGPDIGVNDSSSQLEALLQQLLGRNISLAVATSSPTLATSTDLQVLA 178
Qy 380 HSGYVYKVRKPL--LKAGIKLYELOPNHVAVPATKDKLGTGSSVTSLSHAKTFIVDGKRFIFG 437
Db 179 RGAHVR-QVPMGRLTRGV-----LHSGFWVVDGRHIYMG 211
Qy 438 SFNLDPRSARLNTMGVVI-ESPKIAEQMERT 468
Db 212 SANMDWRSLTQVKELGAVIYNCSHLAQDLEKT 243
```

```
RESULT 14
US-10-242-902-42
; Sequence 42, Application US/10242902
; Publication No. US20060073549A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin
; APPLICANT: Eaton, Dan
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Goddard, Audrey
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin
; APPLICANT: Smith, Victoria
; APPLICANT: Stephan, Jean-Phillippe
; APPLICANT: Watanabe, Colin
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; APPLICANT: Fong, Sherman
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3630R1C54
; CURRENT FILING DATE: 2002-09-13
; PRIOR APPLICATION NUMBER: 10/197942
```



```

, PRIOR FILING DATE: 1997-11-10
, PRIOR APPLICATION NUMBER: 60/079689
, PRIOR FILING DATE: 1998-03-27
, PRIOR APPLICATION NUMBER: 60/086478
, PRIOR FILING DATE: 1998-05-22
, PRIOR APPLICATION NUMBER: 60/087607
, PRIOR FILING DATE: 1998-06-02
, PRIOR APPLICATION NUMBER: 60/089801
, PRIOR FILING DATE: 1998-06-18
, PRIOR APPLICATION NUMBER: 60/090557
, PRIOR FILING DATE: 1998-06-24
, PRIOR APPLICATION NUMBER: 60/090689
, PRIOR FILING DATE: 1998-08-25
, Remaining Prior Application data re
, NUMBER OF SEQ ID NOS: 116
, SEQ ID NO 42
, LENGTH: 489
, TYPE: PRT
, ORGANISM: Homo Sapien
US-10-243-189-42

```

Query Match	3.8%;	Score 104.5;	DB 9;	Length 489;
Best Local Similarity	25.0%;	Pred. No. 1.7;		
Matches	53;	Conservative 33;	Mismatches 71;	Indels 55; Gaps 11;
QY	284	SPYQKIQT-GRIDWOSVOTRLISDPKAGLDRDRKP---	PIAGRLQDALKQPEKSVYL	339
Db	60	SPAWEPLEAEARQORUSCOLVIVESTP-QDLPSAAGSPSAQPIQCAWLQDLDTAQESVHV	118	
QY	340	VSYPFYVPT-----KSGTDALAKLVQ----	DCIDVTVLTN-----LQATD--VAAV	379
Db	119	ASYYSWLTGPDIGWNDSSSGEALLQQLLGRNISLAVATSSPTLARTSTDLQVLAA	178	
QY	380	HSGYVVKRKL--LKAGIKLYELQPNHAVPATKDKLTGSSVTSLSHAKTIFVQGRKIFIG	437	
Db	179	RGAHVR-QVPMGRLTGV-----	LHSKFVWVDGRHIYMG	211
QY	438	SFNLDPRSRARLNTMGVVI-ESPKIAEQMERT	468	
Db	212	SANMDWRSLTQVKELGAVYNCSHLADLEKT	243	

```

RESULT 18
US-10-243-215-42
; Sequence 42, Application US/10243215
; Publication No. US20060073551A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin
; APPLICANT: Eaton, Dan
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Goddard, Audrey
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin
; APPLICANT: Smith, Victoria
; APPLICANT: Stephan, Jean-Philippe
; APPLICANT: Watanabe, Colin
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; APPLICANT: Fong, Sherman
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3630RIC27
; CURRENT APPLICATION NUMBER: US/10/243,215
; CURRENT FILING DATE: 2002-09-12
; PRIOR APPLICATION NUMBER: 10/197942
; PRIOR FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/059114
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/063046
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/065027
; PRIOR FILING DATE: 1997-11-10
; PRIOR APPLICATION NUMBER: 60/079689

```

```
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/086478
; PRIOR FILING DATE: 1998-05-22
; PRIOR APPLICATION NUMBER: 60/087607
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/089801
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/090557
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090689
; PRIOR FILING DATE: 1998-06-25
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 116
; SEQ ID NO 42
; LENGTH: 489
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-243-215-42

Query Match      3.8%; Score 104.5; DB 9; Length 489;
Best Local Similarity 25.0%; Pred. No. 1.7;
Matches 53; Conservative 33; Mismatches 71; Indels 55; Gaps 11;

Qy 284 SPLYQKIQT-GRIDWQSVQTRLISDTPAKGLDRDRKP---PIAGRLQDALKQPEKSVYL 339
Db 60 SPAWEPLAEARQQRDSCQLVLVESIP-QDLPSAAGSPSAQPLGQAWLQLLDTAQESVHV 118
Qy 340 VSPFVFPVT-----KSGTDALAKLVQ---DGIDVTVLTN-----LOATD--VAAV 379
Db 119 ASYYWSLTGPDIGVNDSSQLGEALQKLQQLGRNISLAVATSSPTLARTSTDLQVLAA 178
Qy 380 HSGVVKYRKPL--LKAGIKLYELOPNHAVPATKDKGLTGSSVTSLSHAKTFIVDGKRIFIG 437
Db 179 RGAHVR-QVPMGRLTRGV-----LHSEFWVDGRIHYMG 211
Qy 438 SFNLDPRSARLNTMGVVI-ESPKIAEQMERT 468
Db 212 SANMDWRSLTQVKELGAVIYNCSHLAQDLKRT 243

RESULT 19
US-10-243-236-42
; Sequence 42, Application US/10243236
; Publication No. US20060073552A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin
; APPLICANT: Eaton, Dan
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Goddard, Audrey
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin
; APPLICANT: Smith, Victoria
; APPLICANT: Stephan, Jean-Philippe
; APPLICANT: Watanabe, Colin
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; APPLICANT: Fong, Sherman
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3630R1C35
; CURRENT FILING DATE: 2002-09-12
; PRIOR APPLICATION NUMBER: 10/197942
; PRIOR FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/059114
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/063046
; PRIOR FILING DATE: 1997-11-10
; PRIOR APPLICATION NUMBER: 60/065027
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/079689
; PRIOR FILING DATE: 1997-11-10
; PRIOR APPLICATION NUMBER: 60/086478
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/086478
```

```
; PRIOR FILING DATE: 1998-05-22
; PRIOR APPLICATION NUMBER: 60/087607
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/089801
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/090557
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090689
; PRIOR FILING DATE: 1998-06-25
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 116
; SEQ ID NO 42
; LENGTH: 489
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-243-236-42

Query Match      3.8%; Score 104.5; DB 9; Length 489;
Best Local Similarity 25.0%; Pred. No. 1.7;
Matches 53; Conservative 33; Mismatches 71; Indels 55; Gaps 11;

Qy 284 SPLYQKIQT-GRIDWQSVQTRLISDTPAKGLDRDRKP---PIAGRLQDALKQPEKSVYL 339
Db 60 SPAWEPLAEARQQRDSCQLVLVESIP-QDLPSAAGSPSAQPLGQAWLQLLDTAQESVHV 118
Qy 340 VSPFVFPVT-----KSGTDALAKLVQ---DGIDVTVLTN-----LOATD--VAAV 379
Db 119 ASYYWSLTGPDIGVNDSSQLGEALQKLQQLGRNISLAVATSSPTLARTSTDLQVLAA 178
Qy 380 HSGVVKYRKPL--LKAGIKLYELOPNHAVPATKDKGLTGSSVTSLSHAKTFIVDGKRIFIG 437
Db 179 RGAHVR-QVPMGRLTRGV-----LHSEFWVDGRIHYMG 211
Qy 438 SFNLDPRSARLNTMGVVI-ESPKIAEQMERT 468
Db 212 SANMDWRSLTQVKELGAVIYNCSHLAQDLKRT 243

RESULT 20
US-10-243-298-42
; Sequence 42, Application US/10243298
; Publication No. US20060073553A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin
; APPLICANT: Eaton, Dan
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Goddard, Audrey
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin
; APPLICANT: Smith, Victoria
; APPLICANT: Stephan, Jean-Philippe
; APPLICANT: Watanabe, Colin
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; APPLICANT: Fong, Sherman
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3630R1C49
; CURRENT FILING DATE: 2002-09-13
; PRIOR APPLICATION NUMBER: 10/197942
; PRIOR FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/059114
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/063046
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/065027
; PRIOR FILING DATE: 1997-11-10
; PRIOR APPLICATION NUMBER: 60/079689
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/086478
; PRIOR FILING DATE: 1998-05-22
; PRIOR APPLICATION NUMBER: 60/087607
```

```
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/089801
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/090557
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090689
; PRIOR FILING DATE: 1998-06-25
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 116
; SEQ ID NO 42
; LENGTH: 489
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-243-298-42

Query Match          3.8%; Score 104.5; DB 9; Length 489;
Best Local Similarity 25.0%; Pred. No. 1.7;
Matches 53; Conservative 33; Mismatches 71; Indels 55; Gaps 11;

Qy 284 SPLYQKIQT-GRIDWQSVQTRLSIDTPAKGLDRDRKRP---PIAGRLQDALKQPEKSVYL 339
Db 60 SPAWEPLAEARQORDSCQLVLVESIP-QDLPSAAGSPSAQPLQQLWLLDTAQESVHV 118
Qy 340 VSPYFVPT-----KSGTDALAKLVQ---DGIDVTVLNS-----LQATD--VAAV 379
Db 119 ASYYWSLTGPDIGVNDSSSQLGEALLQQLQGLGRNISLAVATSSPTLARTSTDLQVLAA 178
Qy 380 HSGVVKYRKL--LKAGIKLYELQPNHAVPATKDKGLTGSSVTSLSHAKTFIVDGKRFIG 437
Db 179 RGAHVR-QVPMGRTRGV-----LHSGFWVVDGRHIYMG 211
Qy 438 SFNLDPRSARLNTMGVVI-ESPKIARQMERT 468
Db 212 SANMDWRSLSQVKELGAVIYNCSHLAQDLEKT 243
```

```
RESULT 21
US-10-243-304-42
; Sequence 42, Application US/10243304
; Publication No. US20060073554A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin
; APPLICANT: Eaton, Dan
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Goddard, Audrey
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin
; APPLICANT: Smith, Victoria
; APPLICANT: Stephan, Jean-Philippe
; APPLICANT: Watanabe, Colin
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; APPLICANT: Fong, Sherman
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3630R1C42
; CURRENT APPLICATION NUMBER: US/10/243,304
; CURRENT FILING DATE: 2002-09-12
; PRIOR APPLICATION NUMBER: 10/157942
; PRIOR FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/059114
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/063046
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/065027
; PRIOR FILING DATE: 1997-11-10
; PRIOR APPLICATION NUMBER: 60/079689
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/086478
; PRIOR FILING DATE: 1998-05-22
; PRIOR APPLICATION NUMBER: 60/087607
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/089801
; PRIOR FILING DATE: 1998-06-18
```

```
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/090557
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090689
; PRIOR FILING DATE: 1998-06-25
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 116
; SEQ ID NO 42
; LENGTH: 489
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-243-304-42

Query Match          3.8%; Score 104.5; DB 9; Length 489;
Best Local Similarity 25.0%; Pred. No. 1.7;
Matches 53; Conservative 33; Mismatches 71; Indels 55; Gaps 11;

Qy 284 SPLYQKIQT-GRIDWQSVQTRLSIDTPAKGLDRDRKRP---PIAGRLQDALKQPEKSVYL 339
Db 60 SPAWEPLAEARQORDSCQLVLVESIP-QDLPSAAGSPSAQPLQQLWLLDTAQESVHV 118
Qy 340 VSPYFVPT-----KSGTDALAKLVQ---DGIDVTVLNS-----LQATD--VAAV 379
Db 119 ASYYWSLTGPDIGVNDSSSQLGEALLQQLQGLGRNISLAVATSSPTLARTSTDLQVLAA 178
Qy 380 HSGVVKYRKL--LKAGIKLYELQPNHAVPATKDKGLTGSSVTSLSHAKTFIVDGKRFIG 437
Db 179 RGAHVR-QVPMGRTRGV-----LHSGFWVVDGRHIYMG 211
Qy 438 SFNLDPRSARLNTMGVVI-ESPKIARQMERT 468
Db 212 SANMDWRSLSQVKELGAVIYNCSHLAQDLEKT 243
```

```
RESULT 22
US-10-243-338-42
; Sequence 42, Application US/10243338
; Publication No. US20060073579A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin
; APPLICANT: Eaton, Dan
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Goddard, Audrey
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin
; APPLICANT: Smith, Victoria
; APPLICANT: Stephan, Jean-Philippe
; APPLICANT: Watanabe, Colin
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; APPLICANT: Fong, Sherman
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3630R1C52
; CURRENT APPLICATION NUMBER: US/10/243,338
; CURRENT FILING DATE: 2002-09-13
; PRIOR APPLICATION NUMBER: 10/197942
; PRIOR FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/059114
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/063046
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/065027
; PRIOR FILING DATE: 1997-11-10
; PRIOR APPLICATION NUMBER: 60/079689
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/086478
; PRIOR FILING DATE: 1998-05-22
; PRIOR APPLICATION NUMBER: 60/087607
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/089801
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/090557
```

```
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090689
; PRIOR FILING DATE: 1998-06-25
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 116
; SEQ ID NO 42
; LENGTH: 489
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-243-338-42

Query Match      3.8%; Score 104.5; DB 9; Length 489;
Best Local Similarity 25.0%; Pred. No. 1.7;
Matches 53; Conservative 33; Mismatches 71; Indels 55; Gaps 11;

Qy 284 SPLVQKIQT-GRIDMQSVQTRLSIDTPAKGLDRDRRKP-PIAGRLQDALKQPEKSVYL 339
|||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 60 SPAWEPLAEARQORDSCQLVLVESIP-QDLPAAAGSFAQLGQAWLQLLDTAQESVHV 118
|||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Qy 340 VSPYFVPT-----KSGTDALAKLVQ---DGIDVTVLTVNS----LQATD--VAAV 379
|||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 119 ASYWSLTGPDIGVNDSSQLGEALLQQLLGRNISLAVATSSPTLARTSTDLQVLAA 178
|||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Qy 380 HSGVYKVKRKL--LKAGIKLYELQPNHAVPATKDKGLTGSSVTSLHAKTFIVDGKRIFIG 437
|||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 179 RGAHVR-QVPMGRLTRGV-----LHSEKFWVDGGRHIYMG 211
|||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Qy 438 SFNLDPRSARLNTMGVVI-ESPKIAEQMERT 468
|||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 212 SANMDWRSLTQVKELGAVIYNCSHLAQDLEKT 243
|||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

RESULT 23
US-10-243-345-42
; Sequence 42, Application US/10243345
; Publication No. US20060073555A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin
; APPLICANT: Eaton, Dan
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Goddard, Audrey
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin
; APPLICANT: Smith, Victoria
; APPLICANT: Stephan, Jean-Philippe
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; APPLICANT: Fong, Sherman
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3630R1C33
; CURRENT APPLICATION NUMBER: US/10/243,345
; CURRENT FILING DATE: 2002-09-12
; PRIOR FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/059114
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/063046
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/065027
; PRIOR FILING DATE: 1997-11-10
; PRIOR APPLICATION NUMBER: 60/079689
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/086478
; PRIOR FILING DATE: 1998-05-22
; PRIOR APPLICATION NUMBER: 60/086478
; PRIOR FILING DATE: 1998-05-22
; PRIOR APPLICATION NUMBER: 60/087607
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/089801
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/090557
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/090557
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090689
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090689
; Remaining Prior Application data removed - See File Wrapper or PALM.
```

```
; PRIOR FILING DATE: 1998-06-25
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 116
; SEQ ID NO 42
; LENGTH: 489
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-243-345-42

Query Match      3.8%; Score 104.5; DB 9; Length 489;
Best Local Similarity 25.0%; Pred. No. 1.7;
Matches 53; Conservative 33; Mismatches 71; Indels 55; Gaps 11;

Qy 284 SPLVQKIQT-GRIDMQSVQTRLSIDTPAKGLDRDRRKP-PIAGRLQDALKQPEKSVYL 339
|||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 60 SPAWEPLAEARQORDSCQLVLVESIP-QDLPAAAGSFAQLGQAWLQLLDTAQESVHV 118
|||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Qy 340 VSPYFVPT-----KSGTDALAKLVQ---DGIDVTVLTVNS----LQATD--VAAV 379
|||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 119 ASYWSLTGPDIGVNDSSQLGEALLQQLLGRNISLAVATSSPTLARTSTDLQVLAA 178
|||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Qy 380 HSGVYKVKRKL--LKAGIKLYELQPNHAVPATKDKGLTGSSVTSLHAKTFIVDGKRIFIG 437
|||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 179 RGAHVR-QVPMGRLTRGV-----LHSEKFWVDGGRHIYMG 211
|||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Qy 438 SFNLDPRSARLNTMGVVI-ESPKIAEQMERT 468
|||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 212 SANMDWRSLTQVKELGAVIYNCSHLAQDLEKT 243
|||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

RESULT 24
US-10-243-357-42
; Sequence 42, Application US/10243357
; Publication No. US20060073556A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin
; APPLICANT: Eaton, Dan
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Goddard, Audrey
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin
; APPLICANT: Smith, Victoria
; APPLICANT: Stephan, Jean-Philippe
; APPLICANT: Watanabe, Colin
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; APPLICANT: Fong, Sherman
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3630R1C55
; CURRENT APPLICATION NUMBER: US/10/243,357
; CURRENT FILING DATE: 2002-09-13
; PRIOR FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/059114
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/063046
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/065027
; PRIOR FILING DATE: 1997-11-10
; PRIOR APPLICATION NUMBER: 60/079689
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/086478
; PRIOR FILING DATE: 1998-05-22
; PRIOR APPLICATION NUMBER: 60/087607
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/089801
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/090557
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090689
; PRIOR FILING DATE: 1998-06-25
; Remaining Prior Application data removed - See File Wrapper or PALM.
```



```
; ORGANISM: Homo Sapien
US-10-247-015-42

Query Match          3.8%; Score 104.5; DB 9; Length 489;
Best Local Similarity 25.0%; Pred. No. 1.7;
Matches 53; Conservative 33; Mismatches 71; Indels 55; Gaps 11;

QY 284 SPLYQKIQT-GRIDWQSVQTRLSITPAKGLDRDRKP---PIAGRLQDALKQPEKSVYL 339
||: : : ||: : : ||: : : ||: : : ||: : : ||: : : ||: : : ||: : :
Db 60 SPAWPLEAEARQQRDSCQLVLVESIP-QDLPSAAGSPSAQLGQAWLQLDLTAQESVHV 118

QY 340 VSPYFVPT-----KSGTDALAKLVQ---DGIDVTVLVNS-----LOATD--VAAV 379
||: : : ||: : : ||: : : ||: : : ||: : : ||: : : ||: : : ||: : :
Db 119 ASYWSLTGPDIGVNDSSQLGEALLQKQLLGRNISLAVATSPPTLARTSTDLOVLAA 178

QY 380 HSGYVKYRKPL--LKAGIKLYELOPNHAVPATPKDKGLTGSSVTSLSHAKTFIVDGKRIFIG 437
||: : : ||: : : ||: : : ||: : : ||: : : ||: : : ||: : : ||: : :
Db 179 RCAHVR-QVPMGRLTRGV-----LHSEKFWVDGRHIYWG 211

QY 438 SFNLDPKRSARLNTMGVVI-ESPKIAEQMERT 468
||: : : ||: : : ||: : : ||: : : ||: : : ||: : : ||: : : ||: : :
Db 212 SANMDWRSLTQVKELGAVIYNCSHLAODLEKT 243

RESULT 27
US-11-157-996-2
; Sequence 2, Application US/11157996
; Publication No. US20060003370A1
; GENERAL INFORMATION:
; APPLICANT: BAKER, KEVIN P.
; APPLICANT: CHEN, JIAN
; APPLICANT: DESNOYERS, LUC
; APPLICANT: GODDARD, AUDREY
; APPLICANT: GODOWSKI, PAUL J.
; APPLICANT: GURNEY, AUSTIN L.
; APPLICANT: PAN, JAMES
; APPLICANT: SMITH, VICTORIA
; APPLICANT: WATANABE, COLIN K.
; APPLICANT: WOOD, WILLIAM I.
; APPLICANT: ZHANG, ZEMIN
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES
; FILE REFERENCE: 39780-3630R1C25C1
; CURRENT APPLICATION NUMBER: US/11/157,996
; CURRENT FILING DATE: 2003-06-20
; PRIOR APPLICATION NUMBER: US 10/242,505
; PRIOR FILING DATE: 2002-09-11
; PRIOR APPLICATION NUMBER: US 10/197,942
; PRIOR FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: PCT/US01/27099
; PRIOR FILING DATE: 2001-08-29
; PRIOR APPLICATION NUMBER: PCT/US00/03565
; PRIOR FILING DATE: 2000-02-11
; PRIOR APPLICATION NUMBER: US 60/140,653
; PRIOR FILING DATE: 1999-06-22
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 489
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-157-996-2

Query Match          3.8%; Score 104.5; DB 11; Length 489;
Best Local Similarity 25.0%; Pred. No. 1.7;
Matches 53; Conservative 33; Mismatches 71; Indels 55; Gaps 11;

QY 284 SPLYQKIQT-GRIDWQSVQTRLSITPAKGLDRDRKP---PIAGRLQDALKQPEKSVYL 339
||: : : ||: : : ||: : : ||: : : ||: : : ||: : : ||: : : ||: : :
Db 60 SPAWPLEAEARQQRDSCQLVLVESIP-QDLPSAAGSPSAQLGQAWLQLDLTAQESVHV 118

QY 340 VSPYFVPT-----KSGTDALAKLVQ---DGIDVTVLVNS-----LOATD--VAAV 379
||: : : ||: : : ||: : : ||: : : ||: : : ||: : : ||: : : ||: : :
Db 119 ASYWSLTGPDIGVNDSSQLGEALLQKQLLGRNISLAVATSPPTLARTSTDLOVLAA 178

QY 380 HSGYVKYRKPL--LKAGIKLYELOPNHAVPATPKDKGLTGSSVTSLSHAKTFIVDGKRIFIG 437
||: : : ||: : : ||: : : ||: : : ||: : : ||: : : ||: : : ||: : :
Db 179 RCAHVR-QVPMGRLTRGV-----LHSEKFWVDGRHIYWG 211

QY 438 SFNLDPKRSARLNTMGVVI-ESPKIAEQMERT 468
||: : : ||: : : ||: : : ||: : : ||: : : ||: : : ||: : : ||: : :
Db 212 SANMDWRSLTQVKELGAVIYNCSHLAODLEKT 243

RESULT 28
US-11-226-701-15
; Sequence 15, Application US/11226701
; Publication No. US20060009632A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: Kapeller-Libermann, Rosana
; APPLICANT: Robison, Keith E.
; APPLICANT: White, David
; APPLICANT: Williamson, Mark W.
; APPLICANT: Cook, William James
; APPLICANT: Meyers, Rachel E.
; APPLICANT: MacBeth, Kyle J.
; APPLICANT: Carroll, Joseph M.
; APPLICANT: Chun, Miyoung
; TITLE OF INVENTION: NOVEL 27875, 22025, 27420, 17906, 16319,
; FILE REFERENCE: MPI03-0210MNIM
; CURRENT APPLICATION NUMBER: US/11/226,701
; CURRENT FILING DATE: 2005-09-14
; PRIOR APPLICATION NUMBER: US/10/386,414
; PRIOR FILING DATE: 2003-03-11
; PRIOR APPLICATION NUMBER: 09/426,282
; PRIOR FILING DATE: 1999-10-25
; PRIOR APPLICATION NUMBER: 09/668,266
; PRIOR FILING DATE: 2000-09-22
; PRIOR APPLICATION NUMBER: 09/330,970
; PRIOR FILING DATE: 1999-06-11
; PRIOR APPLICATION NUMBER: 09/724,599
; PRIOR FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: 09/860,193
; PRIOR FILING DATE: 2001-05-16
; PRIOR APPLICATION NUMBER: 09/571,689
; PRIOR FILING DATE: 2000-05-16
; PRIOR APPLICATION NUMBER: 10/283,023
; PRIOR FILING DATE: 2002-10-29
; PRIOR APPLICATION NUMBER: 60/335,044
; PRIOR FILING DATE: 2001-10-31
; PRIOR APPLICATION NUMBER: 10/010,943
; PRIOR FILING DATE: 2001-12-06
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15
; LENGTH: 506
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-11-226-701-15

Query Match          3.8%; Score 104.5; DB 11; Length 506;
Best Local Similarity 25.0%; Pred. No. 1.8;
Matches 53; Conservative 33; Mismatches 71; Indels 55; Gaps 11;

QY 284 SPLYQKIQT-GRIDWQSVQTRLSITPAKGLDRDRKP---PIAGRLQDALKQPEKSVYL 339
||: : : ||: : : ||: : : ||: : : ||: : : ||: : : ||: : : ||: : :
Db 77 SPAWPLEAEARQQRDSCQLVLVESIP-QDLPSAAGSPSAQLGQAWLQLDLTAQESVHV 135

QY 340 VSPYFVPT-----KSGTDALAKLVQ---DGIDVTVLVNS-----LOATD--VAAV 379
||: : : ||: : : ||: : : ||: : : ||: : : ||: : : ||: : : ||: : :
Db 136 ASYWSLTGPDIGVNDSSQLGEALLQKQLLGRNISLAVATSPPTLARTSTDLOVLAA 195

QY 380 HSGYVKYRKPL--LKAGIKLYELOPNHAVPATPKDKGLTGSSVTSLSHAKTFIVDGKRIFIG 437
```


QY 311 KGLDRDRKPPKIAGRLQDALKQPEKSVVLYSPYFVP---TKSGTDALAKLVQDGDIVTVL 367
Db 555 VTEDRD-----ALIDAAADKIEKDLILLGSTAVEDKLGQGVPCIEKLSQAGVKIWL 607
QY 368 TNSLQATDVAHVHSGYVKRKPPLKAGIKLYELOPNHAVPATKDKGLTGSSVTSIHAQTF 427
Db 608 TGDKTET---AINIGYA---CSLLREGMK-----QILVTLT-----SSDIEALEKQG- 648
QY 428 IVDGKRIFIGSFN-----LDRPSARLNTM-GVVIESPKEIAEQWERTLAD 471
Db 649 --DKEAVAKASFQSIKKQLREGMSQTAATAVNTSAKENSEMFLVIDGKSLTYALDSKLEK 706
QY 472 TTPEYAYR 479
Db 707 EFLELAIR 714

RESULT 32
US-11-096-568A-27847
; Sequence 27847, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nickolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 27847
; LENGTH: 1181
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(1181)
; OTHER INFORMATION: Ceres Seq. ID no. 2049239
US-11-096-568A-27847

Query Match 3.7%; Score 101.5; DB 11; Length 1181;
Best Local Similarity 21.2%; Pred. No. 11;
Matches 78; Conservative 61; Mismatches 154; Indels 75; Gaps 18;

QY 151 SHPNIEVRLFPFVLKRWALGYL-TDPPLNRRMHNKSFTA---DNPATILGGRNIGDE 206
Db 492 NQPNAL-----IQKFFRVLAICHTAIPDVNSDTGEITYEASPDAAAFVIASRELGE 545
QY 207 YPKVGEDTV-FADLDILATGVSUVGEVSHDFRYWASHAHNATRIIRSGN-----IGKGL 260
Db 546 FFSRSQTSISLHEIDHM-TGEKVDVRYELLHVLFFSSRKMSVIVRNPNRLLLSKGA 604
QY 261 QALGYNDETSRHALLRYRETVEQSPLYOK--IQTGRIDMQSV-----QTRLI SDTPA 310
Db 605 DSVMPK-RLAKHGQRNERETKEHIKKYAEAGRLTVITYREIDEDEVIVWEEEFNAKTL 663
QY 311 KGLDRDRKPPKIAGRLQDALKQPEKSVVLYSPYFVP---TKSGTDALAKLVQDGDIVTVL 367
Db 664 VTEDRD-----ALIDAAADKIEKDLILLGSTAVEDKLGQGVPCIEKLSQAGVKIWL 716
QY 368 TNSLQATDVAHVHSGYVKRKPPLKAGIKLYELOPNHAVPATKDKGLTGSSVTSIHAQTF 427
Db 717 TGDKTET---AINIGYA---CSLLREGMK-----QILVTLT-----SSDIEALEKQG- 757
QY 428 IVDGKRIFIGSFN-----LDRPSARLNTM-GVVIESPKEIAEQWERTLAD 471
Db 758 --DKEAVAKASFQSIKKQLREGMSQTAATAVNTSAKENSEMFLVIDGKSLTYALDSKLEK 815
QY 472 TTPEYAYR 479
Db 816 EFLELAIR 823

RESULT 33

US-11-096-568A-27846
; Sequence 27846, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nickolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 27846
; LENGTH: 1189
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(1189)
; OTHER INFORMATION: Ceres Seq. ID no. 2049238
US-11-096-568A-27846

Query Match 3.7%; Score 101.5; DB 11; Length 1189;
Best Local Similarity 21.2%; Pred. No. 12;
Matches 78; Conservative 61; Mismatches 154; Indels 75; Gaps 18;

QY 151 SHPNIEVRLFPFVLKRWALGYL-TDPPLNRRMHNKSFTA---DNPATILGGRNIGDE 206
Db 500 NQPNAL-----IQKFFRVLAICHTAIPDVNSDTGEITYEASPDAAAFVIASRELGE 553
QY 207 YPKVGEDTV-FADLDILATGVSUVGEVSHDFRYWASHAHNATRIIRSGN-----IGKGL 260
Db 554 FFSRSQTSISLHEIDHM-TGEKVDVRYELLHVLFFSSRKMSVIVRNPNRLLLSKGA 612
QY 261 QALGYNDETSRHALLRYRETVEQSPLYOK--IQTGRIDMQSV-----QTRLI SDTPA 310
Db 613 DSVMPK-RLAKHGQRNERETKEHIKKYAEAGRLTVITYREIDEDEVIVWEEEFNAKTL 671
QY 311 KGLDRDRKPPKIAGRLQDALKQPEKSVVLYSPYFVP---TKSGTDALAKLVQDGDIVTVL 367
Db 672 VTEDRD-----ALIDAAADKIEKDLILLGSTAVEDKLGQGVPCIEKLSQAGVKIWL 724
QY 368 TNSLQATDVAHVHSGYVKRKPPLKAGIKLYELOPNHAVPATKDKGLTGSSVTSIHAQTF 427
Db 725 TGDKTET---AINIGYA---CSLLREGMK-----QILVTLT-----SSDIEALEKQG- 765
QY 428 IVDGKRIFIGSFN-----LDRPSARLNTM-GVVIESPKEIAEQWERTLAD 471
Db 766 --DKEAVAKASFQSIKKQLREGMSQTAATAVNTSAKENSEMFLVIDGKSLTYALDSKLEK 823
QY 472 TTPEYAYR 479
Db 824 EFLELAIR 831

RESULT 34
US-09-941-095-85
; Sequence 85, Application US/09941095
; Publication No. US20060068383A1
; GENERAL INFORMATION:
; APPLICANT: BROW, MARY ANN D.
; LYAMICHEV, VICTOR I.
; OLIVE, DAVID M.
; TITLE OF INVENTION: RAPID DETECTION AND IDENTIFICATION OF
; PATHOGENS
; NUMBER OF SEQUENCES: 165
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MEDLEN & CARROLL
; STREET: 220 MONTGOMERY STREET, SUITE 2200
; CITY: SAN FRANCISCO
; STATE: CALIFORNIA
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 94104
; COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA: US/09/941,095
FILING DATE: 28-Aug-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/941,193
FILING DATE: 28-Aug-2001
ATTORNEY/AGENT INFORMATION:
NAME: CARROLL, PETER G.
REGISTRATION NUMBER: 32,837
REFERENCE/DOCKET NUMBER: FORS-01756
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338
INFORMATION FOR SEQ ID NO: 85:
SEQUENCE CHARACTERISTICS:
LENGTH: 833 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 85:
US-09-941-095-85

Query Match 3.7%; Score 101; DB 7; Length 833;
Best Local Similarity 21.6%; Pred. No. 7.4;
Matches 118; Conservative 65; Mismatches 175; Indels 188; Gaps 30;
QY 118 NLMYLAERGVV-----RLDDNNTRGL--DLLLLDS---HPNIEVRLNPF 163
DB 321 DLLAALAAAGRVHRAPEPKALRDLKEARGLAKDLVLAIRGLGLPGDDPML---- 376
QY 164 VLKRWALGYLTD----FPLNRRMKNKFT--ADNRATI-----LGGNIGDE---- 206
DB 377 -----LAYLDPSTTPEGVARRYGGEWEAGERAAALSERLFLANLWGRLEGERLLW 429
QY 207 -YFKVGE--DTVPADLILATG-----SVVGEVS-----HDFRYW 239
DB 430 LYREVERPLSAVLAHME--ATGVRLDVAYLRALSLEVAGEIARLEAEVFRLAGHPFN-- 484
QY 240 ASHSAHNATRIIRSGNIGKGLQALGYNDTSRHA-----LLRYRET 281
DB 485 -LNSRDQLERVLFD---ELGLPAIGKTEKGRSTSAVLEALREAHPIVEKILQYREL 540
QY 282 EQ-----SPYQKI--QTGRIDWQSVOT-----RLISDTPAKGLDRDRKPPPIAGRLQDA 329
DB 541 KLKSTYIDPLDPLIHPTTGLTRFNQTATATGRSLSSDP--NLQNIPTVPTLQRIIRA 598
QY 330 LKQEKSVYLSPY-----FVPTKSGTDALAKLVQGDIDVTVLTS-----LQATDVA 377
DB 599 FIAEEGVLVALDYSQIELRLVLAHLSGDNLRVFOEGRDIHTETASWVFQPREAVD-- 656
QY 378 AVHSGYVKYRKPLKA-----GKLYELQPNHVP----- 407
DB 657 -----PLMRAAKTINFCVLYGMSAHLRSQELAIPIYEEAQAFIERYSQSPKVR 705
QY 408 -----ATKDKGLTSSVTSLSHAKTFIVDGKRFIFSGFNLDPRSLRNTEMGVVIESP--K 460
DB 706 AWIEKTLLEGRRRGYVETLF-----GRRYVPDLEARKVSVREAAER--MAFNMPVRGT 757
QY 461 IAEOMERTLADTTP---EYAYRVTLDKHNLQHWHDPATRTYTPNEPEAKLWKRIAKILS 517
DB 758 AADLMKLAVMKFLPRLEMGARMLLQVHDELVLAEAPKRA-----EAVARLAKVMEGVYP 813
QY 518 L-LPIE 522
DB 814 LAVPLE 819

RESULT 35
US-11-198-746-85
Sequence 85, Application US/11/198746
Publication No. US20060035256A1
GENERAL INFORMATION:
APPLICANT: BROW, MARY ANN D.
APPLICANT: LYAMICHEV, VICTOR I.
APPLICANT: OLIVE, DAVID M.
TITLE OF INVENTION: RAPID DETECTION AND IDENTIFICATION OF
TITLE OF INVENTION: PATHOGENS
NUMBER OF SEQUENCES: 165
CORRESPONDENCE ADDRESS:
ADDRESSEE: MEDLEN & CARROLL
STREET: 220 MONTGOMERY STREET, SUITE 2200
CITY: SAN FRANCISCO
STATE: CALIFORNIA
COUNTRY: UNITED STATES OF AMERICA
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/11/198,746
FILING DATE: 05-Aug-2005
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/941,193
FILING DATE: 28-Aug-2001
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: CARROLL, PETER G.
REGISTRATION NUMBER: 32,837
REFERENCE/DOCKET NUMBER: FORS-01756
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338
INFORMATION FOR SEQ ID NO: 85:
SEQUENCE CHARACTERISTICS:
LENGTH: 833 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-11-198-746-85

Query Match 3.7%; Score 101; DB 11; Length 833;
Best Local Similarity 21.6%; Pred. No. 7.4;
Matches 118; Conservative 65; Mismatches 175; Indels 188; Gaps 30;
QY 118 NLMYLAERGVV-----RLDDNNTRGL--DLLLLDS---HPNIEVRLNPF 163
DB 321 DLLAALAAAGRVHRAPEPKALRDLKEARGLAKDLVLAIRGLGLPGDDPML---- 376
QY 164 VLKRWALGYLTD----FPLNRRMKNKFT--ADNRATI-----LGGNIGDE---- 206
DB 377 -----LAYLDPSTTPEGVARRYGGEWEAGERAAALSERLFLANLWGRLEGERLLW 429
QY 207 -YFKVGE--DTVPADLILATG-----SVVGEVS-----HDFRYW 239
DB 430 LYREVERPLSAVLAHME--ATGVRLDVAYLRALSLEVAGEIARLEAEVFRLAGHPFN-- 484
QY 240 ASHSAHNATRIIRSGNIGKGLQALGYNDTSRHA-----LLRYRET 281
DB 485 -LNSRDQLERVLFD---ELGLPAIGKTEKGRSTSAVLEALREAHPIVEKILQYREL 540
QY 282 EQ-----SPYQKI--QTGRIDWQSVOT-----RLISDTPAKGLDRDRKPPPIAGRLQDA 329
DB 541 KLKSTYIDPLDPLIHPTTGLTRFNQTATATGRSLSSDP--NLQNIPTVPTLQRIIRA 598
QY 330 LKQEKSVYLSPY-----FVPTKSGTDALAKLVQGDIDVTVLTS-----LQATDVA 377

Db 599 FIAEGWLLVALDYQIELRVLAHLSGDNELIRVFOEGRDHITETASWFMGVPREAVD-- 656
Qy 378 AVHSGYVYKRPPLKA-----GIKLYELOPNHAVP----- 407
Db 657 -----PLMRAAKTINFGVLYGMSAHLRSQELAIPIYEEAQAFIERVQSFQPKVR 705
Qy 408 ----ATKDKGLTSSVTSIHLAKTFIVDGKRFIFGSLNLDPRSARLNTMGVVIESP---K 460
Db 706 AWIEKTLSEGRRGYVETLF-----GRRRYVPDLERVKSVREAAER--MAFNMPVRGT 757
Qy 461 IAEOMERTLADTTP---EYAVRVTLDKHNRLOQWHDPAIRKTYPNEPEAKLWKRIAANKLS 517
Db 758 AADLMKLVKLPFRLEMGARMLLOVHDELVLPAKERA-----EAVARLAKEVMEGVYP 813
Qy 518 L-LPIE 522
Db 814 LAVPLE 819
RESULT 36
US-11-198-794-85
; Sequence 85, Application US/11198794
; Publication No. US20060035257A1
; GENERAL INFORMATION:
; APPLICANT: BROW, MARY ANN D.
; APPLICANT: LYAMICHEV, VICTOR I.
; TITLE OF INVENTION: RAPID DETECTION AND IDENTIFICATION OF
; TITLE OF INVENTION: PATHOGENS
; NUMBER OF SEQUENCES: 165
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MEDLEN & CARROLL
; STREET: 220 MONTGOMERY STREET, SUITE 2200
; CITY: SAN FRANCISCO
; STATE: CALIFORNIA
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/11/198,794
; FILING DATE: 05-Aug-2005
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/941,193
; FILING DATE: 28-Aug-2001
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: CARROLL, PETER G.
; REGISTRATION NUMBER: 32,837
; REFERENCE/DOCKET NUMBER: FORS-01756
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 705-8410
; TELEFAX: (415) 397-8338
; INFORMATION FOR SEQ ID NO: 85:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 833 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-11-198-794-85
Query Match 3.7%; Score 101; DB 11; Length 833;
Best Local Similarity 21.6%; Pred. No. 7.4;
Matches 118; Conservative 65; Mismatches 175; Indels 188; Gaps 30;

Qy 118 NMVLAERGVVRV-----RLLDDNNTRGI---DDLALDS----HPNIEVRLFNPF 163
Db 321 DLLALAAAGGVRHRAPEPYKALRDLKEARGLAKDLSVLALREGILPPGDDPML---- 376

Qy 164 VLRKWRALGYLTD---FPLNRRMHNKSF--ADNRATI-----LGGNIGDE----- 206
Db 377 -----LAVLLDSNTTPEGVARRYGGEWTEEAGEAALSERLFANLWGRLEGERLLW 429
Qy 207 -YKVGGE--DTVPADLDLILATG-----SVVGEVS-----HDFDRYW 239
Db 430 LYREVERPLSAVLAHME--ATGVRLDVAYLRALSLEVAGEIARLEAEVFRLAGHPFN--- 484
Qy 240 ASHSAHNATRIIRSGNIGKGLQALGYNDTSRHA-----LLRYRETV 281
Db 485 -LNSRDQLERVLFD---ELGLPAIGKTEKTKGRSTSAAVLEALREAHPIVEKILQYRELT 540
Qy 282 EQ-----SPLYQKI--OTGRIDWQSVOT-----RLISDTPAKGLDRDRRKPPIAGRILODA 329
Db 541 KLKSTYIDPLPLDILHPRTGRHLTRFNQTATATGRLSSSDP--NLQNIPIVPTPLQQRIRRA 598
Qy 330 LKQPEKSVYLVSPY-----FVPTKSGTDALAKLVQDGIDVTVLTS-----LQATDVA 377
Db 599 FIAEGWLLVALDYQIELRVLAHLSGDNELIRVFOEGRDHITETASWFMGVPREAVD-- 656
Qy 378 AVHSGYVYKRPPLKA-----GIKLYELOPNHAVP----- 407
Db 657 -----PLMRAAKTINFGVLYGMSAHLRSQELAIPIYEEAQAFIERVQSFQPKVR 705
Qy 408 ----ATKDKGLTSSVTSIHLAKTFIVDGKRFIFGSLNLDPRSARLNTMGVVIESP---K 460
Db 706 AWIEKTLSEGRRGYVETLF-----GRRRYVPDLERVKSVREAAER--MAFNMPVRGT 757
Qy 461 IAEOMERTLADTTP---EYAVRVTLDKHNRLOQWHDPAIRKTYPNEPEAKLWKRIAANKLS 517
Db 758 AADLMKLVKLPFRLEMGARMLLOVHDELVLPAKERA-----EAVARLAKEVMEGVYP 813
Qy 518 L-LPIE 522
Db 814 LAVPLE 819
RESULT 37
US-11-198-657-85
; Sequence 85, Application US/11198657
; Publication No. US20060040299A1
; GENERAL INFORMATION:
; APPLICANT: BROW, MARY ANN D.
; APPLICANT: LYAMICHEV, VICTOR I.
; TITLE OF INVENTION: RAPID DETECTION AND IDENTIFICATION OF
; TITLE OF INVENTION: PATHOGENS
; NUMBER OF SEQUENCES: 165
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MEDLEN & CARROLL
; STREET: 220 MONTGOMERY STREET, SUITE 2200
; CITY: SAN FRANCISCO
; STATE: CALIFORNIA
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/11/198,657
; FILING DATE: 05-Aug-2005
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/941,193
; FILING DATE: 28-Aug-2001
; ATTORNEY/AGENT INFORMATION:
; NAME: CARROLL, PETER G.
; REGISTRATION NUMBER: 32,837
; REFERENCE/DOCKET NUMBER: FORS-01756
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 705-8410
; TELEFAX: (415) 397-8338
; INFORMATION FOR SEQ ID NO: 85:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 833 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-11-198-657-85
Query Match 3.7%; Score 101; DB 11; Length 833;
Best Local Similarity 21.6%; Pred. No. 7.4;
Matches 118; Conservative 65; Mismatches 175; Indels 188; Gaps 30;

```
; TELEPHONE: (415) 705-8410
; TELEFAX: (415) 397-8338
; INFORMATION FOR SEQ ID NO: 85:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 833 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-11-198-657-85

Query Match 3.7%; Score 101; DB 11; Length 833;
Best Local Similarity 21.6%; Pred. No. 7.4;
Matches 118; Conservative 65; Mismatches 175; Indels 188; Gaps 30;

Qy 118 NMVLAARGVVRV-----RLLDNNTRGL---DDLALLDS-----HPNIEVRLNPF 163
Db 321 DLLAARAGVRHAPYKALRDLKEARGLAKLDSVLAIRGLGLPPGDDPML----- 376

Qy 164 VLRKWRALGYLTD---PPRLNRRMHNKSFT--ADNRATI-----LGGRNIGDE--- 206
Db 377 -----LAYLDFSNTTPEGVARRYGGEWTEEAGERAAALSERLFANLWGLEGEERLLW 429

Qy 207 -YFKVGE--DTVPADLDLATG-----SVVGEVS-----HDFRYW 239
Db 430 LYREVERPLSAVLAME--ATGVRLDVAYLRALSLEVAGEIARLEAEVFRLAGHPFN--- 484

Qy 240 ASHAHNATRIIRSGNIGKGLQALGYNDETSRHA-----LRYRETV 281
Db 485 -LNSRDQLERVLFD---ELGUPAIGKTEKGRKSTSAVLEALREAHPIVEKILQYRELT 540

Qy 282 EQ-----SPLYQKI--QTGRIDWQSVOT-----RLISDTPAKGLDRRRKPPPIAGRLQDA 329
Db 541 KLKSTYIDPLDLPRLHPTGRGLHTRFNQTATAGLSSSDP--NLQNTIPVTPQLQRRRA 598

Qy 330 LKQPEKSVYLVSPY-----FVPTKSGTDALAKLVQDGIDVTVLNLS-----LQATDVA 377
Db 599 FIABEGWLLVALDYSQIELRVLAHLSGDENLIRVFQEGRDITHTETASWFMGVPREAVD-- 656

Qy 378 AVHSGYVYKRPPLKA-----GKLYELQPNHNP----- 407
Db 657 -----PLMRRAAKTNFVGLMGSAHRUSQELAIPIYEAQAQFIERYFQSPFKVR 705

Qy 408 -----ATKDKGLTGSVTSLSHAKTFIVDGKRFIFGSFNLDPRSRARLNTMGMVIESP--K 460
Db 706 AWIEKTLTEGRRRGYVETLF-----GRRRYVPDLEARKVSVREAAER-WAFNMPVRGT 757

Qy 461 IAEQWERTLADTTP---EYAYRVTLDKHNRLQWHDPATRKTYPNPEPAKLWKRIAAILKS 517
Db 758 AADLMKLAIVKLPRLBEMGARMLLQVHDELVLPAKERA-----EAVARLAKEVMGVPY 813

Qy 518 L-LPIE 522
Db 814 LAVPLE 819

RESULT 38
US-11-096-568A-30809
; Sequence 30809, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nikolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096.568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 30809
; LENGTH: 1057
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:

Query Match 3.7%; Score 101; DB 11; Length 1166;
Best Local Similarity 21.3%; Pred. No. 12;
Matches 76; Conservative 60; Mismatches 153; Indels 68; Gaps 18;

Qy 151 SHPNIEVRLNPFVLRKWRALGYL-TDFPRLNRRMHNKSFTA---DNRATILGGRNIGDE 206
Db 492 NQFNAEL-----IQKFFRVLAICHTAIPDVNSDTGEITYEASPDFAAFVIASRELGF 545

Qy 207 YFKVGEDTV-FADLDILATGVSVVGEVSHDFRYWASHSAHNATRIIRSGN-----IQKGL 260
Db 546 FFSRSQTSISLHEIDHM---TVYELLHVLE---FSSSRKRMVIVRNPNRLLLSKGA 598

Qy 261 QALGYNDETSRHALLRYRETVESPLYQK--IQTGRIDWQSV-----QTRLISDTPA 310
Db 599 DSVNFK-RLAKHGRQNERETKEHIKKYAEAGLRTLVTITYREIDEDVYVWEEEFNAKTL 657

Qy 311 KGLDRRRKPPPIAGRLQDALKQPEKSVYLVSPYFVP---TKSGTDALAKLVQDGIDVTVL 367
```

```
; NAME/KEY: misc feature
; LOCATION: (1)-(11057)
; OTHER INFORMATION: Ceres Seq. ID no. 4973364
US-11-096-568A-30809

Query Match 3.7%; Score 101; DB 11; Length 1057;
Best Local Similarity 21.3%; Pred. No. 11;
Matches 76; Conservative 60; Mismatches 153; Indels 68; Gaps 18;

Qy 151 SHPNIEVRLNPFVLRKWRALGYL-TDFPRLNRRMHNKSFTA---DNRATILGGRNIGDE 206
Db 383 NQFNAEL-----IQKFFRVLAICHTAIPDVNSDTGEITYEASPDFAAFVIASRELGF 436

Qy 207 YFKVGEDTV-FADLDILATGVSVVGEVSHDFRYWASHSAHNATRIIRSGN-----IQKGL 260
Db 437 FFSRSQTSISLHEIDHM---TVYELLHVLE---FSSSRKRMVIVRNPNRLLLSKGA 489

Qy 261 QALGYNDETSRHALLRYRETVESPLYQK--IQTGRIDWQSV-----QTRLISDTPA 310
Db 490 DSVNFK-RLAKHGRQNERETKEHIKKYAEAGLRTLVTITYREIDEDVYVWEEEFNAKTL 548

Qy 311 KGLDRRRKPPPIAGRLQDALKQPEKSVYLVSPYFVP---TKSGTDALAKLVQDGIDVTVL 367
Db 549 VTEDRD-----ALIDAAAADKIEKDLILGSTAVEDKLGKVPDCIEKLSQAGVKIWWL 601

Qy 368 TNSLQATDVAHVSGYVYKRPPLKAGIK--LYELQPN--HAVPATKDKGLTGSVTSLSH 423
Db 602 TGGKTET--AINIGVA---CSLLREGMKQILVTLDSDDIEALEKQGDKEAVAKLRGMS 655

Qy 424 AKTFIVDGKRFIFGSFNLDPRSRARLNTM-GVVIESPKEIAEQMERTLADTTPEYAYR 479
Db 656 QTAAVTD-----NSAKENSEMFLVDGKSLTYALDLSKLEKEFLELAIR 699

RESULT 39
US-11-096-568A-30808
; Sequence 30808, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nikolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096.568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 30808
; LENGTH: 1166
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:

Query Match 3.7%; Score 101; DB 11; Length 1166;
Best Local Similarity 21.3%; Pred. No. 12;
Matches 76; Conservative 60; Mismatches 153; Indels 68; Gaps 18;

Qy 151 SHPNIEVRLNPFVLRKWRALGYL-TDFPRLNRRMHNKSFTA---DNRATILGGRNIGDE 206
Db 492 NQFNAEL-----IQKFFRVLAICHTAIPDVNSDTGEITYEASPDFAAFVIASRELGF 545

Qy 207 YFKVGEDTV-FADLDILATGVSVVGEVSHDFRYWASHSAHNATRIIRSGN-----IQKGL 260
Db 546 FFSRSQTSISLHEIDHM---TVYELLHVLE---FSSSRKRMVIVRNPNRLLLSKGA 598

Qy 261 QALGYNDETSRHALLRYRETVESPLYQK--IQTGRIDWQSV-----QTRLISDTPA 310
Db 599 DSVNFK-RLAKHGRQNERETKEHIKKYAEAGLRTLVTITYREIDEDVYVWEEEFNAKTL 657

Qy 311 KGLDRRRKPPPIAGRLQDALKQPEKSVYLVSPYFVP---TKSGTDALAKLVQDGIDVTVL 367
```

```

|||||
658 VTEDRD-----ALIDAAADKIEKDLILLGSTAVEDKLQGVPCIEKLSQAGVKIWWL 710
QY  TNSLQATDVAVHSGYVKYRKPPLLKAGIK--LYELQPN--HAVPATKDKGLTSSVTSLSH 423
Db  TCDKTET---AINIGYA---CSLLREGMKQILVTLDDSSDIEALEKQGDKEAVAKLREGMS 764
QY  AKTFIVDGKRIFIGSFNLDPRSARLNTM-GVWIESPKIAEQMERTLADTTPEYAYR 479
Db  QTAAVTD-----NSAKENSEMFLGVIDGKSLTYALDSKLEKEFLELAIR 808

RESULT 40
US-11-096-568A-30807
; Sequence 30807, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nickolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; TITLE OF INVENTION: Therby
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 30807
; LENGTH: 1174
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(1174)
; OTHER INFORMATION: Ceres Seq. ID no. 4973362
US-11-096-568A-30807

Query Match          3.7%; Score 101; DB 11; Length 1174;
Best Local Similarity 21.3%; Pred No. 12;
Matches 76; Conservative 60; Mismatches 153; Indels 68; Gaps 18;

QY  151 SHPNIEVRLNFPVLVRKWRALGYL-TDPPRLNRRMHNKSFTA---DNRATILGGRNIGDE 206
Db  500 NQNAEL-----IQKFFRVLAICHTAIPDVNSDTGEITYEAESPDEAAAFVIASRELGFE 553
QY  207 YKVGEDTV-FADLDILATGVSUVGVSHDFDRYWASHSAHNATRIIRSGN-----IGKGL 260
Db  554 FFSRSQTSISLHEIDHM----TVYELHLVLE---FSSSRKRMSVIVRNPNENRLLLSKGA 606
QY  261 QALGYNDETSRHALLRYRETVESQPLYOK--IQTGRIDWQSV-----QTRLISDTPA 310
Db  607 DSVMEFK-PLAKHGRQNERETKEHIKKAEGAGLRTIVITYREIDEDYEIVWEEEFNAKTL 665
QY  311 KGLDRDRRKPPIAGRLQDALQKPEKSVYLVSPYFVP---TKSGTDALAKLYQDGDIVTVL 367
Db  666 VTEDRD-----ALIDAAADKIEKDLILLGSTAVEDKLQGVPCIEKLSQAGVKIWWL 718
QY  368 TNSLQATDVAVHSGYVKYRKPPLLKAGIK--LYELQPN--HAVPATKDKGLTSSVTSLSH 423
Db  719 TGDKTET---AINIGYA---CSLLREGMKQILVTLDDSSDIEALEKQGDKEAVAKLREGMS 772
QY  424 AKTFIVDGKRIFIGSFNLDPRSARLNTM-GVWIESPKIAEQMERTLADTTPEYAYR 479
Db  773 QTAAVTD-----NSAKENSEMFLGVIDGKSLTYALDSKLEKEFLELAIR 816
```

Search completed: May 2, 2006, 06:28:54
Job time : 31 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: May 2, 2006, 06:06:20 ; Search time 232 Seconds
(without alignments)
1596.562 Million cell updates/sec

Title: US-10-665-990A-14
Perfect score: 2720
Sequence: 1 MHTDPKIQAMPSETISPMKT.....KMKRIAATLSLPIEGLL 525

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Uniprot_05.80.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2720	100.0	525	2 Q6W764_NEIGO	Q6W764 neisseria g
2	2712	99.7	525	2 Q5F887_NEIG1	Q5F887 neisseria g
3	2643	97.2	525	2 Q9JTT1_NEIMA	Q9JTT1 neisseria m
4	2573	94.6	508	2 Q9JYU0_NEIME	Q9JYU0 neisseria m
5	1182	43.5	524	2 Q7VUJ6_BORPE	Q7VUJ6 bordetella
6	1181	43.4	492	2 Q7WPE3_BORBR	Q7WPE3 bordetella
7	1104.5	40.6	493	2 Q8FIS6_ECOL6	Q8FIS6 escherichia
8	1102.5	40.5	493	1 YMDC_ECOL1	P75919 escherichia
9	1102.5	40.5	493	2 Q83RUS_SHIFL	Q83RUS shigella fl
10	1100.5	40.5	493	2 Q7AF48_ECO57	Q7AF48 escherichia
11	1098	40.4	528	2 Q57QL0_SALCH	Q57QL0 salmonella
12	1093.5	40.2	493	2 Q8X917_ECO57	Q8X917 escherichia
13	1092.5	40.2	514	2 Q9CKM5_PASMU	Q9CKM5 pasteurella
14	1089.5	40.1	494	2 Q8ZQ28_SALTY	Q8ZQ28 salmonella
15	1089.5	40.1	495	2 Q8Z7M0_SALTI	Q8Z7M0 salmonella
16	1089.5	40.1	495	2 Q5PGV9_SALPA	Q5PGV9 salmonella
17	940.5	34.6	519	2 Q6F8A6_AC1AD	Q6F8A6 acinetobact
18	881.5	32.4	516	2 Q4LR64_9BURK	Q4LR64 burkholderi
19	881	32.4	505	2 Q87JB2_VIBPA	Q87JB2 vibrio para
20	872.5	32.1	501	2 Q8D641_VIBVU	Q8D641 vibrio vuln
21	871.5	32.0	524	2 Q8D861_VIBVU	Q8D861 vibrio vuln
22	869.5	32.0	554	2 Q7MMB3_VIBVU	Q7MMB3 vibrio vuln
23	862.5	31.7	501	2 Q8D800_VIBVU	Q8D800 vibrio vuln
24	850.5	31.3	520	2 Q8P1M4_XANAC	Q8P1M4 xanthomonas
25	849.5	31.2	530	2 Q6MQ17_EDEAM	Q6MQ17 bdellovibri
26	848.5	31.2	570	2 Q4FTL4_9GAMM	Q4FTL4 psychrobact
27	846	31.1	464	2 Q5QUY8_IDILO	Q5QUY8 idiomarina
28	843	31.0	520	2 Q4UWU7_XANCP	Q4UWU7 xanthomonas
29	843	31.0	520	2 Q8P7A5_XANCP	Q8P7A5 xanthomonas
30	828	30.4	523	2 Q6G185_BAROU	Q6G185 bartonella
31	825	30.3	523	2 Q6G4X0_BARHE	Q6G4X0 bartonella

RESULT 1

ID	Q6W764_NEIGO	466	2	Q98N23_RHILO	Q98N23 rhizobium l
AC	Q6W764;	518	2	Q8UEX3_AGRTS	Q8UEX3 agrobacteri
DT	05-JUL-2004 (TrEMBLrel. 27, Created)	466	2	Q984B3_RHILO	Q984B3 rhizobium l
DT	05-JUL-2004 (TrEMBLrel. 27, Last sequence update)	520	2	Q6FB53_AC1AD	Q6FB53 acinetobact
DT	05-JUL-2004 (TrEMBLrel. 27, Last annotation update)	529	2	Q9HTP4_PSEAE	Q9HTP4 pseudomonas
DE	Phospholipase D (Fragment).	529	2	Q62E43_BURMA	Q62E43 burkholderi
GN	Name-pld;	517	2	Q88CA5_PSEPK	Q88CA5 pseudomonas
OS	Neisseria gonorrhoeae.	540	2	Q4LJD9_9BURK	Q4LJD9 burkholderi
OC	Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;	550	2	Q63J33_BURPS	Q63J33 burkholderi
OC	Neisseriaceae; Neisseria.	510	2	Q6LI13_PHOPR	Q6LI13 photobacter
OX	NCBI_TaxID=485;	542	2	Q88BC2_PSESM	Q88BC2 pseudomonas
RN	[1]	526	2	Q4ZZX0_PSESY	Q4ZZX0 pseudomonas
RP	NUCLEOTIDE SEQUENCE.	515	2	Q5E616_VIBF1	Q5E616 vibrio fisc
RC	STRAN=1291;	521	2	Q92Q03_RHIME	Q92Q03 rhizobium m
RX	MEDLINE=22935225; PubMed=14573659;				
RR	DOI=10.1128/IAI.71.11.6381-6391.2003;				
RA	Edwards J.L., Entz D.D., Apicella M.A.;				
RT	"Gonococcal phospholipase d modulates the expression and function of complement receptor 3 in primary cervical epithelial cells.";				
RL	Infect. Immun. 71:6381-6391(2003).				
DR	EMBL; AY307929; AAQ77232.1; -; Genomic DNA.				
DR	GO; GO:0003824; F:catalytic activity; IEA.				
DR	GO; GO:0008152; P:metabolism; IEA.				
DR	InterPro; IPR001736; PLD.				
DR	Pfam; PF00614; PLDC; 2.				
DR	SMART; SM00155; PLDC; 2.				
DR	PROSITE; PS00035; PLD; 2.				
FT	NON TER 525 525				
SQ	SEQUENCE 525 AA; 59224 MW; 4B626F5264E076D1 CRC64;				
Query Match 100.0%; Score 2720; DB 2; Length 525;					
Best Local Similarity 100.0%; Pred. No. 7.6e-184;					
Matches 525; Conservative 0; Mismatches 0; Indels 0; Gaps 0;					
Qy	1	MHTDPKIQAMPSETISPMKTRSLISLCLLCCSSWLPPLERTSRHNTSKPVLLDN	60		
Db	1	MHTDPKIQAMPSETISPMKTRSLISLCLLCCSSWLPPLERTSRHNTSKPVLLDN	60		
Qy	61	ILQIRHTPHNNGSLDIYLLDDPHEALAAALIESAEHSLDQYIYWRNDISGRLLFNL	120		
Db	61	ILQIRHTPHNNGSLDIYLLDDPHEALAAALIESAEHSLDQYIYWRNDISGRLLFNL	120		
Qy	121	YLAERGVRLLLDDNNTRGLDILLALDSHNIEVRLNPFVLRKWRALGYLTDPPRL	180		
Db	121	YLAERGVRLLLDDNNTRGLDILLALDSHNIEVRLNPFVLRKWRALGYLTDPPRL	180		
Qy	181	NRRMHNKSFADNRATILGGRNIGDEYFKVGEDTVFADLDILATGSVGVSHDFDRYWA	240		
Db	181	NRRMHNKSFADNRATILGGRNIGDEYFKVGEDTVFADLDILATGSVGVSHDFDRYWA	240		
Qy	241	SHSAHWATRIIRSGNTGKGLQALQYNDTSRHALRYRETVEQSPLYQKIQTCRIDMQSV	300		
Db	241	SHSAHWATRIIRSGNTGKGLQALQYNDTSRHALRYRETVEQSPLYQKIQTCRIDMQSV	300		

ALIGNMENTS

```
Db 241 SHSAHNATRIIRSGNIGKGLQALQVNDTSHALLRYRETVEQSPLYQKIQTGRIDWQSV 300
Qy 301 QTRLSIDTPAKGLDRDRKPPPIAGRLQDALQKPEKSVYLVSPFYVPTKSGTDLAKLVQD 360
Db 301 QTRLSIDTPAKGLDRDRKPPPIAGRLQDALQKPEKSVYLVSPFYVPTKSGTDLAKLVQD 360
Qy 361 GIDVTVLNSLOATDVAHVHSGYVKYRKLKAGIKLYELQPNHAVPATKDKGLTGSSVT 420
Db 361 GIDVTVLNSLOATDVAHVHSGYVKYRKLKAGIKLYELQPNHAVPATKDKGLTGSSVT 420
Qy 421 SLHAKTFTVDGKRIFIGSFNLDPRSARLNTMGVVIESPKIAEQMERTLADTTPEYAYRV 480
Db 421 SLHAKTFTVDGKRIFIGSFNLDPRSARLNTMGVVIESPKIAEQMERTLADTTPEYAYRV 480
Qy 481 TLDKHNRLQWHDHPATRKTYPNEPEAKLWKRIAANKILSLPIEGLL 525
Db 481 TLDKHNRLQWHDHPATRKTYPNEPEAKLWKRIAANKILSLPIEGLL 525

RESULT 2
Q5F887 NEIG1
ID Q5F887 NEIG1 PRELIMINARY; PRT; 525 AA.
AC Q5F887
DT 10-MAY-2005 (TREMBlrel. 30, Created)
DT 10-MAY-2005 (TREMBlrel. 30, Last sequence update)
DE Putative phospholipase D-family protein.
GN OrderedLocusNames=NGO0902;
OS Neisseria gonorrhoeae (strain ATCC 700825 / FA 1090).
OC Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC Neisseriaceae; Neisseria.
OX NCBI_TaxID=242231;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RA Lewis L.A., Gillaspay A.F., McLaughlin R.E., Gipson M., Ducey T.F.,
RA Ownbey T., Hartman K., Nydick C., Carson M.B., Vaughn J., Thomson C.,
RA Song L., Lin S., Yuan X., Najjar F., Zhan M., Ren Q., Zhu H., Qi S.,
RA Kenton S.M., Lai H., White J.D., Clifton S., Roe B.A., Dyer D.W.;
RA "The complete genome sequence of Neisseria gonorrhoeae.";
RT Submitted (MAR-2003) to the EMBL/GenBank/DDbj databases.
DR EMBL: AE004969; AAW89600.1; -; Genomic_DNA.
DR InterPro: IPR001736; PLD.
DR Pfam: PF00614; PLDC; 2.
DR SMART: SM00155; PLDC; 2.
DR PROSITE: PS50035; PLD; 2.
KW Complete proteome.
SQ SEQUENCE 525 AA; 59244 MW; 67C126AFCDAA0A360 CRC64;

Query Match 99.7%; Score 2712; DB 2; Length 525;
Best Local Similarity 99.6%; Pred. No. 2.8e-183;
Matches 523; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MHTDPKIQAMPSETISPMKTRSLISLLCLLCCSSWLPPLPEERTESRHNTSKPVLADN 60
Db 1 MHTDPKIQAMPSETISPMKTRSLISLLCLLCCSSWLPPLPEERTESRHNTSKPVLADN 60
Qy 61 ILQIRHTPHNGLSDIYLLDDPHEAALAAARALIESAHSLDLQYYIWRNDISGRLLFNLM 120
Db 61 ILQIRHTPHNGLSDIYLLDDPHEAFAARAALIESAHSLDLQYYIWRNDISGRLLFNLM 120
Qy 121 YLAERGVVRVRLDDNNTRGLDLLLLALDHPNIEVLFNPFVLRKWRALGYLTDFFRL 180
Db 121 YLAERGVVRVRLDDNNTRGLDLLLLALDHPNIEVLFNPFVLRKWRALGYLTDFFRL 180
Qy 181 NRRMHNKFTADNRATILGGRNIGDEYFKVGEDTVFADLDILATGSGVGVSHDPRYWA 240
Db 181 NRRMHNKFTADNRATILGGRNIGDEYFKVGEDTVFADLDILATGSGVGVSHDPRYWA 240
Qy 241 SHSAHNATRIIRSGNIGKGLQALQVNDTSHALLRYRETVEQSPLYQKIQTGRIDWQSV 300
Db 241 SHSAHNATRIIRSGNIGKGLQALQVNDTSHALLRYRETVEQSPLYQKIQTGRIDWQSV 300
Qy 301 QTRLSIDTPAKGLDRDRKPPPIAGRLQDALQKPEKSVYLVSPFYVPTKSGTDLAKLVQD 360
```

```
Db 301 QTRLSIDTPAKGLDRDRKPPPIAGRLQDALQKPEKSVYLVSPFYVPTKSGTDLAKLVQD 360
Qy 361 GIDVTVLNSLOATDVAHVHSGYVKYRKLKAGIKLYELQPNHAVPATKDKGLTGSSVT 420
Db 361 GIDVTVLNSLOATDVAHVHSGYVKYRKLKAGIKLYELQPNHAVPATKDKGLTGSSVT 420
Qy 421 SLHAKTFTVDGKRIFIGSFNLDPRSARLNTMGVVIESPKIAEQMERTLADTTPEYAYRV 480
Db 421 SLHAKTFTVDGKRIFIGSFNLDPRSARLNTMGVVIESPKIAEQMERTLADTTPEYAYRV 480
Qy 481 TLDKHNRLQWHDHPATRKTYPNEPEAKLWKRIAANKILSLPIEGLL 525
Db 481 TLDKHNRLQWHDHPATRKTYPNEPEAKLWKRIAANKILSLPIEGLL 525

RESULT 3
Q9JTT1 NEIMA
ID Q9JTT1 NEIMA PRELIMINARY; PRT; 525 AA.
AC Q9JTT1;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE Phospholipase D-family protein.
GN OrderedLocusNames=NMA1646;
OS Neisseria meningitidis (serogroup A).
OC Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC Neisseriaceae; Neisseria.
OX NCBI_TaxID=65699;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Z2491 / Serogroup A / Serotype 4A;
RX MEDLINE=20222556; PubMed=10761919; DOI=10.1038/35006655;
RA Parkhill J., Achtman M., James K.D., Bentley S.D., Churcher C.M.,
RA Klee S.R., Morelli G., Basham D., Brown D., Chillingworth T.,
RA Davies R.M., Davis P., Devlin K., Feltwell T., Hamlin N., Holroyd S.,
RA Jagels K., Leather S., Moule S., Mungall K.L., Quail M.A.,
RA Rajandream M.A., Rutherford K.M., Simmonds M., Skelton J.,
RA Whitehead S., Spratt B.G., Barrrell B.G.;
RT "Complete DNA sequence of a serogroup A strain of Neisseria meningitidis Z2491.";
RL Nature 404:502-506(2000).
DR EMBL: AL162756; CAB84874.1; -; Genomic_DNA.
DR PIR: B81859; B81859.
DR GO: 0003824; F: catalytic activity; IEA.
DR GO: 0008152; P: metabolism; IEA.
DR InterPro: IPR001736; PLD.
DR Pfam: PF00614; PLDC; 2.
DR SMART: SM00155; PLDC; 2.
DR PROSITE: PS50035; PLD; 2.
KW Complete proteome.
SQ SEQUENCE 525 AA; 59305 MW; 1CFB7AC5B82F1B02 CRC64;

Query Match 97.2%; Score 2643; DB 2; Length 525;
Best Local Similarity 97.3%; Pred. No. 2.1e-178;
Matches 511; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

Qy 1 MHTDPKIQAMPSETISPMKTRSLISLLCLLCCSSWLPPLPEERTESRHNTSKPVLADN 60
Db 1 MRPNPKIQAMPSETISLMKTRSLISLLCLLCCSSWLPPLPEERTESRHNTSKPVLADN 60
Qy 61 ILQIRHTPHNGLSDIYLLDDPHEAALAAARALIESAHSLDLQYYIWRNDISGRLLFNLM 120
Db 61 ILQIRHTPHNGLSDIYLLNDPHEAFAARAALIESAHSLDLQYYIWRNDISGRLLFNLM 120
Qy 121 YLAERGVVRVRLDDNNTRGLDLLLLALDHPNIEVLFNPFVLRKWRALGYLTDFFRL 180
Db 121 YLAERGVVRVRLDDNNTRGLDLLLLALDHPNIEVLFNPFVLRKWRALGYLTDFFRL 180
Qy 181 NRRMHNKFTADNRATILGGRNIGDEYFKVGEDTVFADLDILATGSGVGVSHDPRYWA 240
Db 181 NRRMHNKFTADNRATILGGRNIGDEYFKVGEDTVFADLDILATGSGVGVSHDPRYWA 240
```


QY 241 SHSAHNATRIIRSGNIGKGLQALGYNDTSRHALRYRETVEQSPYQKIOTGRIDMQSV 300
 Db 241 SHSAHNATRIIRSGNIGKGLQALGYNDTSRHALRYRETVEQSPYQKIOTGRIDMQSV 300
 QY 301 QTRLISDTPAKGLDRBKPIAGRLQALQKPEKSVYLVSPYFVPTKSGTDALAKLVQD 360
 Db 301 QTRLISDDPAKGLDRBKPIAGRLQALQKPEKSVYLVSPYFVPTKSGTDALAKLVQD 360
 QY 361 GIDVTVLNSLOATDVAHVSGYVKYKRPPLKAGIKLYELOPNHAPVATKDKGLTGSSTV 420
 Db 361 GIDVTVLNSLOATDVAHVSGYVKYKRPPLKAGIKLYELOPNHAPVATKDKGLTGSSTV 420
 QY 421 SLHAKTFIVDGKRFIFGIFGNLDPRLNTMGVVIESPKEAEQWERTLADTTPEYAYRV 480
 Db 421 SLHAKTFIVDGKRFIFGIFGNLDPRLNTMGVVIESPKEAEQWERTLADTTPEYAYRV 480
 QY 481 TLDKHNRLQWHDPAKTRKTYNPEPAKLWKRIAAILSLPIEGLL 525
 Db 481 TLDKHNRLQWHDPAKTRKTYNPEPAKLWKRIAAILSLPIEGLL 525

RESULT 4

Q9JYU0_NEIMB
 ID Q9JYU0_NEIMB PRELIMINARY; PRT; 508 AA.
 AC Q9JYU0;
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Cardiolipin synthetase family protein.
 GN OrderedLocusNames=NM1434;
 OS Neisseria meningitidis (serogroup B).
 OC Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
 OC Neisseriaceae; Neisseria.
 OC NCBI_TaxID=491;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=MC58 / Serogroup B;
 RX MEDLINE=20175755; PubMed=10710307; DOI=10.1126/science.287.5459.1809;
 RA Tettelin H., Saunders N.J., Heidelberg J.F., Jeffries A.C.,
 RA Nelson K.E., Eisen J.A., Ketchum K.A., Hood D.W., Peden J.F.,
 RA Dodson R.J., Nelson W.C., Gwinn M.L., DeBoy R.T., Peterson J.D.,
 RA Hickey E.K., Haft D.H., Salzberg S.L., White O., Fleischmann R.D.,
 RA Dougherty B.A., Mason T.M., Ciecko A., Parksey D.S., Blair E.,
 RA Cittone H., Clark E.B., Cotton M.D., Utterback T.R., Khouri H.M.,
 RA Qin H., Vamathevan J.J., Gill J., Scarlato V., Maignani V., Pizza M.,
 RA Grandi G., Sun L., Smith H.O., Fraser C.M., Moxon E.R., Rappuoli R.,
 RA Venter J.C.;
 RT "Complete genome sequence of Neisseria meningitidis serogroup B strain
 MC58.";
 RL Science 287:1809-1815(2000).
 DR EMBL; AE002098; AAF41795.1; -; Genomic_DNA.
 DR FIR; B81083; B81083.
 DR TIGR; NM1434; -.
 DR GO; GO:0003824; F:catalytic activity; IEA.
 DR GO; GO:0008152; P:metabolism; IEA.
 DR InterPro; IPR001736; PLD.
 DR Pfam; PF00614; PLDC; 2.
 DR SMART; SM00155; PLDC; 2.
 DR PROSITE; PS50035; PLD; 2.
 KW Complete proteome.
 SQ SEQUENCE 508 AA; 57350 MW; A2DE6CAC47CA25D5 CRC64;

Query Match 94.68; Score 2573; DB 2; Length 508;
 Best Local Similarity 97.8; Pred. No. 1.8e-173;
 Matches 497; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

QY 18 MKTRSLISLLCLLSCSSWLPPLTEERTSRHFNTPKVLNLIQIRHTPHNGLSDIY 77
 Db 1 MKTRSLISLLCLLSCSSWLPPLTEERTSRHFNTPKVLNLIQIRHTPHNGLSDIY 60

QY 78 LLDDPHALAAALIESAHSLDQYIWRNDISGRLLFNLMYLAARGVRVRLLLDDN 137
 Db 61 LLNDPHEFAAARAALIESAHSLDQYIWRNDISGRLLFNLMYLAARGVRVRLLLDDN 120

QY 138 NTRGLDLLLALDSHPNIEVLEFNPVLRKWRALGYLTDFPRLNRMHMKSFADNRATI 197
 Db 121 NTRGLDLLLALDSHPNIEVLEFNPVLRKWRALGYLTDFPRLNRMHMKSFADNRATI 180
 QY 198 LGGNIGDEYFKVGEDTVFADLDILATGCVVGEVSHDFDRYWASHSAHNATRIIRSGNIG 257
 Db 181 LGGNIGDEYFKVGEDTVFADLDILATGCVVGEVSHDFDRYWASHSAHNATRIIRSGDIG 240
 QY 258 KGLQALGYNDTSRHALRYRETVEQSPYQKIOTGRIDMQSVQTRLISDTPAKGLDRDR 317
 Db 241 KGLQALGYNDTSRHALRYRETVEQSPYQKIOTGRIDMQSVQTRLISDTPAKGLDRDR 300
 QY 318 RKPIAGRLQALQKPEKSVYLVSPYFVPTKSGTDALAKLVQDGDVTVLTNSLOATDVA 377
 Db 301 RKPIAGRLQALQKPEKSVYLVSPYFVPTKSGTDALAKLVQDGDVTVLTNSLOATDVA 360
 QY 378 AVHSGYVKYKRPPLKAGIKLYELOPNHAPVATKDKGLTGSSTVSLHAKTFIVDGKRFIFG 437
 Db 361 AVHSGYVKYKRPPLKAGIKLYELOPNHAPVATKDKGLTGSSTVSLHAKTFIVDGKRFIFG 420
 QY 438 SFNLDPRLNTMGVVIESPKEAEQWERTLADTTPEYAYRVTLDKHNRLQWHDPAKTRK 497
 Db 421 SFNLDPRLNTMGVVIESPKEAEQWERTLADTTPEYAYRVTLDRHRLQWHDPAKTRK 480
 QY 498 TYPNEPEAKLWKRIAAILSLPIEGLL 525
 Db 481 TYPNEPEAKLWKRIAAILSLPIEGLL 508

RESULT 5

Q7VUJ6_BORPE
 ID Q7VUJ6_BORPE PRELIMINARY; PRT; 524 AA.
 AC Q7VUJ6;
 DT 01-OCT-2003 (TrEMBLrel. 25, Created)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Putative phospholipase D protein.
 GN OrderedLocusNames=BP3092;
 OS Bordetella pertussis.
 OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
 OC Alcaligenaceae; Bordetella.
 OC NCBI_TaxID=520;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=Tohama I / ATCC BAA-589 / NCTC 13251;
 RX MEDLINE=22827954; PubMed=12910271; DOI=10.1038/ng1227;
 RA Parkhill J., Sebatia M., Preston A., Murphy L.D., Thomson N.R.,
 RA Harris D.E., Holden M.T.G., Churcher C.M., Bentley S.D., Mungall K.L.,
 RA Cerdeno-Tarraga A.-M., Temple L., James K.D., Harris B., Quail M.A.,
 RA Achtman M., Atkin R., Baker S., Basham D., Bason N., Cherevach I.,
 RA Chillingworth T., Collins M., Cronin A., Davis P., Doggett J.,
 RA Feltham T., Goble A., Hamlin N., Hauser H., Holroyd S., Jagels K.,
 RA Leather S., Moulé S., Norberczak H., O'Neill S., Ormond D., Price C.,
 RA Rabinowitsch E., Rutter S., Sanders M., Saunders D., Seeger K.,
 RA Sharp S., Simmonds M., Skelton J., Squares R., Squares S., Stevens K.,
 RA Unwin L., Whitehead S., Barrell B.G., Maskell D.J.;
 RT "Comparative analysis of the genome sequences of Bordetella pertussis,
 RT Bordetella parapertussis and Bordetella bronchiseptica.";
 RL Nat. Genet. 35:32-40(2003).
 DR EMBL; BX640420; CAE43360.1; -; Genomic DNA.
 DR GO; GO:0003824; F:catalytic activity; IEA.
 DR GO; GO:0008152; P:metabolism; IEA.
 DR InterPro; IPR001736; PLD.
 DR Pfam; PF00614; PLDC; 2.
 DR SMART; SM00155; PLDC; 2.
 DR PROSITE; PS50035; PLD; 2.
 KW Complete proteome.
 SQ SEQUENCE 524 AA; 56590 MW; EDD16DE78F59B466 CRC64;

Query Match 43.5%; Score 1182; DB 2; Length 524;
 Best Local Similarity 48.6; Pred. No. 5.8e-75;
 Matches 258; Conservative 77; Mismatches 152; Indels 44; Gaps 10;

```
Qy 19 KTRSLISL--LCLLLCSCSSWLPLEERTESRHFNTSKPVLDDNLIQIRHTP----- 68
Db 14 RAALLAAGCAILLGACT--LPPVDRTAGHALDAD-----QARATPLGGQVGE 62
Qy 69 --HNNGLSDIYLLDDPHEALAAALIBSAHSLDLOYYIWRNDISGRLLFNLMYLAAR 126
Db 63 ADTHPLSGFHVLDGAQDAFAARMLLAARATLDDVOYYIWRNDWTGTLLOALHAAER 122
Qy 127 GVRVRLDDNNRGLDLLLLALASHPNIEVRLNPFVLRKWRALGYLTDPRLNRRMHN 186
Db 123 GVRVRLDDNNGISGLDDALAALDAHNAEVRLENFPFTRSFKALGYLTDPSRLNRRMHN 182
Qy 187 KSFTADNRATILIGRNIGDEYFKVGEDTVFADLLIATGSSVGVSHDFDRYWASHAHN 246
Db 183 KSFTVDNQATIIGRNIGDEYFGATDGVLFADLDVAVGVVDVSAEFDAYWASESAMP 242
Qy 247 ATRIIRSGNIGKGLQALGYNDTSRHH-----ALLRYRETVQSPLYQKIOTGRIDWQSVQ 301
Db 243 AGPLLLPA-----PGAQTLRALAERARIEQDPAAGDYSALRELPIRELMAGRPLQWAP 298
Qy 302 TRLISDTPAKGLDRDRRKPIAG----RLQDALKQPEKSVLVSPYFVPTKSGTDALAKL 357
Db 299 ARWSDDDPAKGLG---KAPPAQLLTQQLRNILGEPRTLDLVSPYFVPAEAGTQAFAL 354
Qy 358 VQDGIDVTVLTNSLOATDVAAVHSGYVYKRPKLLKAGIKLYELOPNHAVPATKGLTGS 417
Db 355 ARGGAQVRVLTNALEATDVAVVHSGYAKRRKALLQAGVRLYEMRRSYAGPKQGRFGS 414
Qy 418 SVTSLHAKTFTVDGKRIFGSNLPPRSANLNTMGVVIESPKEIAEQMERTLADTTPEYA 477
Db 415 SGSSLHAKTFGVDGERVIFGSPNPRSANLNTLGFVIESPDMAHIAAATFDQDIPAT 474
Qy 478 YRVLTLDKHNRLQW---HDPATKTYPNPEAKLMKRIAAKILSLPIEGLL 525
Db 475 YEVRLLDDGSLYWLQORDGATVR-HDSEPGVSLWRRFVWLFSLPLPLEPL 524

RESULT 6
Q7WP63 BORBR PRELIMINARY; PRT; 492 AA.
AC Q7WP63;
DT 01-OCT-2003 (TReMBLrel. 25, Created)
DT 01-OCT-2003 (TReMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)
DE Putative phospholipase D protein.
GN OrderedLocusNames=BB0822;
OS Bordetella bronchiseptica (Alcaligenes bronchisepticus).
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Alcaligenaceae; Bordetella.
OX NCBI_TaxID=518;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=RB50 / ATCC BAA-588;
RX MEDLINE=22827954; PubMed=12910271; DOI=10.1038/ngi1227;
RA Parkhill J., Sebahia M., Preston A., Murphy L.D., Thomson N.R.,
RA Harris D.E., Holden M.T.G., Churcher C.M., Bentley S.D., Mungall K.L.,
RA Cerdeno-Tarraga A.-M., Temple L., James K.D., Harris B., Quail M.A.,
RA Achtman M., Atkin R., Baker S., Basham D., Bason N., Cherevach I.,
RA Chillingworth T., Collins M., Cronin A., Davis P., Doggett J.,
RA Feltwell T., Goble A., Hamlin N., Hauser H., Holroyd S., Jagels K.,
RA Leather S., Moule S., Norberczak H., O'Neil S., Ormond D., Price C.,
RA Rabinowitsch E., Rutter S., Sanders M., Saunders D., Seeger K.,
RA Sharp S., Simmonds M., Skelton J., Squares R., Squares S., Stevens K.,
RA Unwin L., Whitehead S., Barrrell B.G., Maskell D.J.
RT "Comparative analysis of the genome sequences of Bordetella pertussis,
RT Bordetella parapertussis and Bordetella bronchiseptica."
RL Nat. Genet. 35:32-40(2003).
DR EMBL; BX640439; CAE31321.1; -; Genomic DNA.
DR GO; GO:0003824; F:metabolic activity; IEA.
DR GO; GO:0008152; F:metabolism; IEA.
DR InterPro; IPR001736; PLD.
DR Pfam; PF00614; PLDc; 2.
```

```
DR SMART; SMO0155; PLDc; 2.
DR PROSITE; PS50035; PLD; 2.
KW Complete proteome.
SQ SEQUENCE 492 AA; 53309 MW; BC6BFEBE8AC34E726 CRC64;

Query Match 43.4%; Score 1181; DB 2; Length 492;
Best Local Similarity 49.6%; Pred. No. 6.3e-75;
Matches 253; Conservative 70; Mismatches 147; Indels 40; Gaps 8;

Qy 38 LPPLLEERTESRHFNTSKPVLDDNLIQIRHTPHNNGLSDI-----YLLDDPHEALA 87
Db 1 MPPVPDRTAGHALDAS-----QARATPLGGQIGDLADAHPLGSGFHALGDAQDAFA 51
Qy 88 ARAALIESAHSILDQYYIWRNDISGRLLFNLMYLAARGVVRVRLDDNNRGLDLLLL 147
Db 52 ARMLAARATRLDVOYYIWRNDWTGTLLOALHAAERGVVRVRLDDNNGISGLDALA 111
Qy 148 ALDSHPNIEVRLNPFVLRKWRALGYLTDPRLNRRMHNKSFADNRATILIGRNIGDEY 207
Db 112 ALDAHNAEVRLENFPFTRSFKALGYLTDPSRLNRRMHNKSFVDNQATIIGRNIGDEY 171
Qy 208 FKVGEDTVFADLLIATGSSVGVSHDFDRYWASHAHNATRIIRSGNIGKGLQALGYND 267
Db 172 FGATDGVLFADLDVAVGVVDVSAEFDAYWASESAMPAGPLLP-----PGAQTLRALA 227
Qy 268 ETSRH-----ALLRYRETVQSPLYQKIOTGRIDWQSVQTRLISDTPAKGLDRDRRKPI 322
Db 228 ERAARIQDDPAAGDYSALRELPIRELMAGRPLQWAPARWSDDDPAKGLG----KAPP 283
Qy 323 AG-----RLQDALKQPEKSVLVSPYFVPTKSGTDALAKLVQDGDIVTVLTNSLOATDVAA 378
Db 284 AGLLTQQLRNILGEPRTLDLVSPYFVPTTEAGTQAFALARGGAQVRVLTNALEATDVAV 343
Qy 379 VHSQYVYKRPKLLKAGIKLYELOPNHAVPATKGLTGSSVTSLHAKTFTVDGKRIFGS 438
Db 344 VHSQYAKRRKALLQAGVRLYEMRRSYAGPKQGRFGSGSSSLHAKTFTVDGGERVIFGS 403
Qy 439 FNLDPRSANLNTMGVVIESPKEIAEQMERTLADTTPEYAVRVLTLDKHNRLQW---HDPAT 495
Db 404 FNFDPRSANLNTLGFVIESPDMAHIAAATFDQDIPATYEVRLDDGSLYWLQORDGAT 463
Qy 496 RKTYPNPEAKLMKRIAAKILSLPIEGLL 525
Db 464 VR-HDSEPGVSLWRRFVWLFSLPLPLEPL 492

RESULT 7
Q8FIS6 ECOL6 PRELIMINARY; PRT; 493 AA.
AC Q8FIS6;
DT 01-MAR-2003 (TReMBLrel. 23, Created)
DT 01-MAR-2003 (TReMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
DE Hypothetical protein ymdC.
GN Name=ymdC; OrderedLocusNames=c1310;
OS Escherichia coli O6
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=217992;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=O6.H1 / CFT073 / ATCC 700928;
RX MEDLINE=22398234; PubMed=12471157; DOI=10.1073/pnas.252529799;
RA Welch R.A., Burland V., Plunkett G. III, Redford P., Roesch P.,
RA Rayko D., Buckles E.L., Liou S.-R., Boutin A., Hackett J., Stroud D.,
RA Mayhew G.F., Rose D.J., Zhou S., Schwartz D.C., Perna N.T.,
RA Mobley H.L.T., Donnenberg M.S., Blattner F.R.;
RT "Extensive mosaic structure revealed by the complete genome sequence
RT of uropathogenic Escherichia coli."
RL Proc. Natl. Acad. Sci. U.S.A. 99:17020-17024 (2002).
DR EMBL; AE016759; AAN7983.1; -; Genomic DNA.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
```

```
DR GO; GO:0008152; P:metabolism; IEA.
DR GO; GO:000654; P:phospholipid biosynthesis; IEA.
DR InterPro; IPR001736; PLD.
DR Pfam; PF00614; PLDC; 2.
DR SMART; SM00155; PLDC; 2.
DR PROSITE; PS00035; PLD; 2.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 493 AA; 55876 MW; 582F673DC3BDC7A CRC64;

Query Match 40.6%; Score 1104.5; DB 2; Length 493;
Best Local Similarity 47.5%; Pred. No. 1.6e-69;
Matches 220; Conservative 85; Mismatches 143; Indels 15; Gaps 5;

Qy 72 GLSDIYLLDDPHEALARAALIESAHSIDLOQYIWRNDISGRLLFNMLYLAERGVRV 131
Db 37 GQCGLPFLSKSLDAFAARYLAEMAEHTLDVQYIWDMSGRLLFSALLAAAKRGVRV 96
Qy 132 LLDDNNTRGLDLLLLALDHPNIEVLFNPFVLRKWRALGYLTDPRLNRRMHNKSF 191
Db 97 LLDDNNTPGLDNLRLDLSHPRIEVLNPFPSFLLRPLGYITDFSLNRRMHNKSF 156
Qy 192 DNRATILGGRNIGDEYFKVGEDTVFADLDILATGSGVVGESHDFDRYASHSAHNATRII 251
Db 157 DGVVTLVGGRNIGDAYFGAGEEPLFSDLDVMAIGPVVEDVADDFARYWYCKSVPLQOVL 216
Qy 252 RSGNICKG-----LQALGYNDSTSRHALLRVRETVESQPLYQKIQTGTIDWQSVQTRL 304
Db 217 ---DVPEGEMADRIELPASWHDAMTH---RYLRKMESSPFINHLVDGTLIWAQTRL 269
Qy 305 ISDTPAKGLDRDRKPPKIAGRLQDALKQPEKSVLVSPYFVPTKSGTDALAKLVQGDIDV 364
Db 270 LSDDPKAGSGKAKRSHLLPQRLFDIMGSPSERDIISSVFPTRAGVAQLLRMVRKGVKI 329
Qy 365 TVLTNSLOATDVAHVSGYVKRPLKAGIKLYELOPNHAVPAT-KDKGLTGSSVTSLSH 423
Db 330 AILTNSLAANDVAHVAGYARWRKLLRYGVELYELKPTREQSSTLHNRGITSNGASLH 389
Qy 424 AKTFIVDGRKIFIGSNLDRPSARLNTMGVIVESPKEIAEQWERTLADTTPAYRVTLDK 483
Db 390 AKTFSIDGKTVFGSNFDPRLTNTMGVIVESLTLAQIDKRFIQSOYDAWQRLD 449
Qy 484 KHNRLQWHDPATRK--TYPNEPEAKLWKRIAAKILSLPIEGLL 525
Db 450 RWGRINWVDRHSKEIILKKEPATSFWRKRVVWKLASILPVEWLL 493

RESULT 8
YMDC_ECOLI STANDARD; PRT; 493 AA.
AC P75919;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Hypothetical protein ymdc.
GN Names:ymdc; OrderedLocusNames=b1046;
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503; DOI=10.1126/science.277.5331.1453;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Berna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12.";
RL Science 277:1453-1474 (1997).
[2]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=K12;
RX MEDLINE=97061202; PubMed=8905232;
[3]
```



```
Db 390 AKTFSIDGKTVFGSFNDFPRSTLLNTMGFVIESETLAQLDKRFTQSYDAAWQLRLD 449
QY 484 KHNRLQWHDPAKTR--TYPNEPEAKLWKRIAAKILSLPIEGLL 525
Db 450 RWGRINWVDRHAKKEIVLKKEPATSFWKRVVMVRLASILPVEWLL 493

RESULT 11
Q57QJ0_SALCH
ID Q57QJ0_SALCH PRELIMINARY; PRT; 528 AA.
AC Q57QJ0;
DT 10-MAY-2005 (TRENBLrel. 30, Created)
DT 10-MAY-2005 (TRENBLrel. 30, Last sequence update)
DT 10-MAY-2005 (TRENBLrel. 30, Last annotation update)
DE Putative phospholipase.
GN Name=yndC; OrderedLocusNames=SC1095;
OS Salmonella cholerae-suis (Salmonella enterica).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
OX NCBI_TaxID=591;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=SC-B67;
RC PubMed=15781495;
RA Chiu C.-H., Tang P., Chu C., Hu S., Bao Q., Yu J., Chou Y.-Y.,
RA Wang H.-S., Lee Y.-S.;
RT "The genome sequence of Salmonella enterica serovar Choleraesuis, a
RT highly invasive and resistant zoonotic pathogen.";
RL Nucleic Acids Res. 33:1690-1698 (2005).
DR EMBL; AE017220; AAK65001.1; -; Genomic_DNA.
KW Complete proteome.
SQ SEQUENCE 528 AA; 60029 MW; D5175EE93D71CDBB CRC64;

Query Match 40.4%; Score 1098; DB 2; Length 528;
Best Local Similarity 44.2%; Pred. No. 5.2e-69;
Matches 235; Conservative 88; Mismatches 175; Indels 34; Gaps 8;

QY 17 PMKTRS-----LISLCLLLCSCSWLPPLEERTESRHNFTSKPVLDDNILQIRHT--- 67
Db 8 PLKSPSGRFOILLPVTLCLNRYTSLVMVKLPGFTQD--YLLSKATTLDPKTRLERAVEP 65
QY 68 --PHNGLSDIYLLDDPHEALARAALIESAHSLLDQYIWRNDISGRLLFNLMYLAEE 125
Db 66 LCARHPGECGIALDNLSDAFAARYLTETAARTLDVQYITWDDMSGRLLFSVLSAAK 125
QY 126 RGVVRLLDDNNTRGLDDLLALDSDHPNIEVLFNFPVLRKWCALGYLTDFFRLNRMH 185
Db 126 RGVHVRLLDDNNTPGLDDTLRLDSDHPNIEVLFNFPVLRKWCALGYLTDFFRLNRMH 185
QY 186 NKSPADNRATILGGRNIGDEYFKVGEDTVFADLDLATGSGVGVSHDFDRYWASHSAH 245
Db 186 NKSYTAGVTVLGVGRNIGDAYFGAGEEPLFSDLVMAIGPVNDVANDERYWRCSSVS 245
QY 246 NATRIIRSGNIGKGLQAL-----GYNDETSRHALLRYRETVEQSPLYQKIQTGRID 296
Db 246 TLQOVLSE-----QELTQRIELPESWYNDIETR----RYLHKLET'SQFMADLDRGRLP 296
QY 297 WQSVOTRLISDTPAKGLDRDRKPPPIAGRLQDALKEKSVLYVSPVFTKSGTDALAK 356
Db 297 LIWAKTRLLSDPSKSGKAQRHSLLPORLFDVNGSPSTERIDIISAYFVTRAGVQLLN 356
QY 357 LVQDGDIVTVLTNSLOATDVAHVSGVYKRPKLLKAGIKYELQPNHA--VPATKDKGLT 415
Db 357 LVKRGVXIALTNSLAANDVAVHAGYARWRKLLRYGVLEYELKPTREHETAVHGRGLT 416
QY 416 GSSVTSLSHAKTIFVDGKRIFIGSNLDPRLSRARLNTMGVGVIESPKIAEQMERTLADTPE 475
Db 417 GNSGSLRHAKTFSIDGSKVFGSLNDFPRSTLLNTMGFVIESETLATLHKHFTQSQRD 476
QY 476 YAYRVTLDKNRLQWHD--PATKTYTNEPEAKLWKRIAAKILSLPIEGLL 525
Db 477 AAQWLRLDRGRINWIDROQEEKEVKKEPATSFWRQVRLVRLAAILPVEWLL 528
```

```
RESULT 12
Q8X9I7_ECO57
ID Q8X9I7_ECO57 PRELIMINARY; PRT; 493 AA.
AC Q8X9I7;
DT 01-MAR-2002 (TRENBLrel. 20, Created)
DT 01-MAR-2002 (TRENBLrel. 20, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE Putative synthase.
GN Name=yndC; OrderedLocusNames=z1680;
OS Escherichia coli O157:H7.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=83334;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=O157:H7 / EDL933 / ATCC 700927 / EHEC;
RC MEDLINE=12074933; PubMed=11206551; DOI=10.1038/35054089;
RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
RA Grotbeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamousis K.,
RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
RA Welch R.A., Blattner F.R.;
RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7.";
RL Nature 409:529-533 (2001).
DR EMBL; AE005174; AAG55792.1; -; Genomic_DNA.
DR PIR; D85666; D85666.
DR PIR; H90806; H90806.
DR GO; GO:0003824; F:catalytic activity; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR001736; PLD.
DR Pfam; PF00614; PLDC; 2.
DR SMART; SM00155; PLDC; 2.
DR PROSITE; PS00035; PLD; 2.
KW Complete proteome.
SQ SEQUENCE 493 AA; 55981 MW; DCC21D8AAE643457 CRC64;

Query Match 40.2%; Score 1093.5; DB 2; Length 493;
Best Local Similarity 47.2%; Pred. No. 9.8e-69;
Matches 219; Conservative 84; Mismatches 144; Indels 17; Gaps 5;

QY 72 GLSDIYLLDDPHEALARAALIESAHSLLDQYIWRNDISGRLLFNLMYLAABRGVVR 131
Db 37 GCGCLFPLEKSLDAFAARYLAEMSEHTLDVQYIWDQDMSGRLLFSALLAAAKRGVVR 96
QY 132 LLLDDNNTRGLDILLALDSDHPNIEVLFNFPVLRKWRALGYLTDFFRLNRMHNSFTA 191
Db 97 LLLDDNNTPGLDDTLRLDSDHPRIEVLFPNFPFRLRLPLGYITDFSLNRRMHNSFTV 156
QY 192 DNRATILGGRNIGDEYFKVGEDTVFADLDLATGSGVGVSHDFDRYWASHSAHNATRII 251
Db 157 DGVTVLVGGRNIGDAYFGAGEEPLFSDLVMAIGPVVEDVADDFARYWYCKSVPLQOVL 216
QY 252 RSGNIGKG-----LQALGYNDETSRHALLRYRETVEQSPLYQKIQTGRIDMQSVOTRL 304
Db 217 ---DVPEGEMADRIEPLASWHDAMTH----RYLRKMESSFNNHLVDGTLPLIWAIRL 269
QY 305 ISDTPAKGLDRDRKPPPIAGRLQDALKEKSVLYVSPVFTKSGTDALAKLVQDGDIV 364
Db 270 LSDDPAKGEKAKRHSLLPQRLFDINGSPSPERIDIISYFVPTPAGVQLLRVMVRKVKI 329
QY 365 TVLTNSLOATDVAHVSGVYKRPKLLKAGIKYELQPNHAPAT--KDKGLTGSSTVSLH 423
Db 330 AILTNSLAANDVAVHAGYARWRKLLRYGVLEYELKPTREQSSTLHDRGITGNSGASLH 389
QY 424 AKTIFVDGKRIFIGSNLDPRLSRARLNTMGVGVIESPKIAEQMERTLADTTPYAYRVTL 483
Db 390 AKTFSIDGKTVFGSLNDFPRSTLLNTMGFVIESETLAQLDKRFTQSYDAAWQLRLD 449
QY 484 KHNRLQWHDPAKTR--TYPNEPEAKLWKRIAAKILSLPIEGLL 525
Db 450 RWGRINWVDRHAKKEIVLKKEPATSFWKRVVMVRLASILPVEWLL 493
```

```
RESULT 13
Q9CKM5_PASMU PRELIMINARY; PRT; 514 AA.
ID Q9CKM5;
AC Q8ZQ28;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein PM1586.
GN OrderedLocusNames=PM1586;
OS Pasteurella multocida.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Pasteurella.
OX NCBI_TaxID=747;
RN NUCLEOTIDE SEQUENCE.
RP STRAIN=Pm70;
RC MEDLINE=21145866; PubMed=11248100; DOI=10.1073/pnas.051634598;
RA May B.J., Zhang Q., Li L.L., Paustian M.L., Whittam T.S., Kapur V.;
RT "Complete genomic sequence of Pasteurella multocida Pm70." ;
RL Proc. Natl. Acad. Sci. U.S.A. 98:3460-3465(2001).
DR EMBL; AE006196; AAK03670.1; -; Genomic DNA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR001736; PLD.
DR Pfam; PF00614; PLDC; 2.
DR SMART; SM00155; PLDC; 2.
DR PROSITE; PS00035; PLD; 2.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 514 AA; 57882 MW; C65B0EB65D80F216 CRC64;

Query Match 40.2%; Score 1092.5; DB 2; Length 514;
Best Local Similarity 46.5%; Pred. No. 1.2e-68;
Matches 238; Conservative 78; Mismatches 173; Indels 23; Gaps 8;

QY 23 LLSLCLLLCCSSLPPLLEERTESRHFTSKPVLNLIQIRH-----TPHNGLSDIY 77
DB 17 LFSVLIALISYQRL--PTSDRPISHT-----PPLNPNGLLARHLLPQVAQHPN-LTGLY 68

QY 78 LLDPPHEALAAARALIESAEHSLSLDIQYIWRNDISGRLLFNLMYLAERGVVRVRLLDN 137
DB 69 PLDGKDAFLARLALSEAEHTLDLQYIYIWHNDVSHLLQSLYKAAVRGVKVRLLLDN 128

QY 138 NTRGLDLLLLALDSDHPNIEVRLFNFPVLRKWRALGYLDFPPLNRRMHNKSFADNRATI 197
DB 129 NTKGMDTTLASINAHPIQIRLNFPMQRYRWLGFLSDFFPLNRRMHNKSFADGVMSI 188

QY 198 LGGRNIGDEYFKVGEDTVFADLDILATGVSVGEVSHDFDRYWASHSAHNATRIIRSGNIG 257
DB 189 LGGRNIGDEYFDVGVGVLFADLDVAATGAVTHIQTDFFRYWNSPSSYPLESIIIRDPI 248

QY 258 KGLQALGYNDETSRRHALLRYRETVBQSPLYQKIOTGRIDWOSVOTRLISDTPAKGLDRDR 317
DB 249 PNPPLALDDEYQ-----TYLKQLTELFPKSLKAGTLAFTWAEALISDDPKKALGKSL 303

QY 318 RKPTIAGRLQDALKQPEKSVLYSPYVPTKSGTDLAKLVDGDDVTVLNSLOATDVA 377
DB 304 TQDSVLAIAPTMLNAKNLIIVSYFVPTTHVGVDFLSRISQGTQGVTSILNLSLEATDVS 363

QY 378 AVHSGVVKYRKPLLKAGIKLYELOPNHAPVAPKDKG--LTGSSVTSLSAKTFIVDGKRF 435
DB 364 IVHSGYAKHRTLLQKQQLYELKP-HATIQMESSGHLKKGASSASLAKFTFLDNRYLF 422

QY 436 IGSFNLDPRLSARLNTMGVWIESPKIAEQMERTLADTTPEYAYRVTLDKHNRLQWHPAT 495
DB 423 VGSFNMDPRSAMUNTEMGLLIDSPELARLLSDGLQONQANYAFSVKLNEAQAALYWEYQEN 482

QY 496 RK--TYPNEPEAKLWKRIAAILLPIEGLL 525
DB 483 GKWIYENEPHTSWFKFVSVCWCLSWLVEHLL 514
```

```
RESULT 14
Q8ZQ28_SALTY PRELIMINARY; PRT; 494 AA.
ID Q8ZQ28;
AC Q8ZQ28;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Putative phospholipase.
GN Name=yndC; OrderedLocusNames=STM1148;
OS Salmonella typhimurium.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
OX NCBI_TaxID=602;
RN NUCLEOTIDE SEQUENCE.
RP STRAIN=LT2;
RC MEDLINE=21534948; PubMed=11677609; DOI=10.1038/35101614;
RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
RA Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,
RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
RA Ryan E., Sun H., Flores L., Miller W., Stoneking T., Nhan M.,
RA Waterston K., Wilson R.K.;
RT "Complete genome sequence of Salmonella enterica serovar Typhimurium LT2." ;
RL Nature 413:852-856(2001).
DR EMBL; AE008749; AAL20078.1; -; Genomic DNA.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR GO; GO:0008654; P:phospholipid biosynthesis; IEA.
DR InterPro; IPR001736; PLD.
DR Pfam; PF00614; PLDC; 2.
DR SMART; SM00155; PLDC; 2.
DR PROSITE; PS00035; PLD; 2.
KW Complete proteome.
SQ SEQUENCE 494 AA; 56217 MW; 571C01EFC4FCB288 CRC64;

Query Match 40.1%; Score 1089.5; DB 2; Length 494;
Best Local Similarity 47.4%; Pred. No. 1.9e-68;
Matches 221; Conservative 77; Mismatches 147; Indels 21; Gaps 5;

QY 72 GLSDIYLLDDPHEALAAARALIESAEHSLSLDIQYIWRNDISGRLLFNLMYLAERGVVR 131
DB 38 GEGCLALDNLDAFAARYRLTEMAARLTDVQYIWEEDMSGRLLFSVLLSAAKGVHVR 97

QY 132 LLLDDNTRGLDLLLLALDSDHPNIEVRLFNFPVLRKWRALGYLDFPPLNRRMHNKSF 191
DB 98 LLLDDNTPGLDDTLRLDSDHPNIEVRLFNFPFRTLRALGYLTDFAELNRRMHNKSYTA 157

QY 192 DNRATILGGRNIGDEYFKVGEDTVFADLDILATGVSVGEVSHDFDRYWASHSAHNATRII 251
DB 158 DGVTVLGGRNIGDAYFGAGEEPLFSDLDVWAIQGVNDVANDFERYWRCSSVSTLQOVL 217

QY 252 RSGNIGKGLQAL-----GYNDETSRRHALLRYRETVEQSPLYQKIOTGRIDWOSVOT 302
DB 218 SLSE-----QELTQRIELPESWYNDIETR----RYLHKLETSSQFMADLDGRGLPLIWA 268

QY 303 RLISDTPAKGLDRDRRPPPIAGRLQDALKQPEKSVLYSPYVPTKSGTDLAKLQVQGI 362
DB 269 RLLSDDPKSGKGAQRHSLLPQRLFDVWMSPTERIDIISAYFVPTPAGVQLLNLRKGV 328

QY 363 DVTVLNLSQATDVAHVHSGVVKYRKPLLKAGIKLYELOPNHA-VPATKDKGLTGSSVTS 421
DB 329 KAILTNLSAANDVAHVHAGYARWRKLLRYGVELYELKPTREHETAVHDLRGLTNGSGSS 388

QY 422 LHAFTFIVDGKRIIFIGSFNLDPRSLNTEMGVWIESPKIAEQMERTLADTTPEYAYRV 481
DB 399 LHAFTFSDGSKVFIGSLNLFDPRLTNTENGFIETSETLATLIHKRFTQSORDAAWQLR 448

QY 482 LDKENRLQWHD--PATRKYTYPNEPEAKLWKRIAAILLPIEGLL 525
DB 449 LDRWGRINWIDRQOEKVKKEPATREFWRQVRLVRLAAILPVEWLL 494
```



```

RESULT 15
Q827M0_SALTI
ID Q827M0_SALTI PRELIMINARY; PRT; 495 AA.
AC Q827M0; OTC9A7;
DT 01-MAR-2002 (TRENBLrel. 20, Created)
DT 01-MAR-2002 (TRENBLrel. 20, Last sequence update)
DT 01-FEB-2005 (TRENBLrel. 29, Last annotation update)
DE Hypothetical protein STY1185.
GN OrderedLocusNames=STY1185, t1772;
OS Salmonella typhi.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
OX NCBI_TaxID=601;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=CT18;
RX MEDLINE=21534947; PubMed=11677608; DOI=10.1038/35101607;
RA Parkhill J., Dougan G., James K.D., Thomson N.R., Pickard D., Wain J.,
RA Churcher C.M., Mungall K.L., Bentley S.D., Holden M.T.G., Sebahia M.,
RA Baker S., Basham D., Brooks K., Chillingworth T., Connerton P.,
RA Cronin A., Davis P., Davies R.M., Dowd L., White N., Farrar J.,
RA Feltwell T., Hamlin N., Haque A., Hien T.T., Holroyd S., Jagels K.,
RA Krogh A., Larsen T.S., Leather S., Moule S., O'Gaora P., Parry C.,
RA Quail M.A., Rutherford K.M., Simmonds M., Skelton J., Stevens K.,
RA Whitehead S., Barrrell B.G.;
RT "Complete genome sequence of a multiple drug resistant Salmonella
RT enterica serovar Typhi CT18."
RL Nature 413:848-852(2001).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=TY2 / ATCC 700931;
RX MEDLINE=22531367; PubMed=12644504;
RX DOI=10.1128/JB.185.7.2330-2337.2003;
RA Deng W., Liou S.-R., Plunkett G. III, Mayhew G.F., Rose D.J.,
RA Burland V., Kodoyanni V., Schwartz D.C., Blattner F.R.;
RT "Comparative genomics of Salmonella enterica serovar Typhi strains Ty2
RT and CT18."
RL J. Bacteriol. 185:2330-2337(2003).
DR EMBL; AL627269; CAD08272.1; -; Genomic DNA.
DR EMBL; AE016840; AO69395.1; -; Genomic DNA.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR GO; GO:0008654; P:phospholipid biosynthesis; IEA.
DR InterPro; IPR001736; PLD.
DR Pfam; PF00614; PLDC; 2.
DR SMART; SM00155; PLDC; 2.
DR PROSITE; PS50035; PLD; 2.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 495 AA; 56348 MW; 4555257D8DDB3AB9 CRC64;

Query Match 40.1%; Score 1089.5; DB 2; Length 495;
Best Local Similarity 47.4%; Pred. No. 1.9e-68;
Matches 221; Conservative 77; Mismatches 147; Indels 21; Gaps 5;

QY 72 GLSDIYLLDDPHEALARAALIESAHSLDQYIWRNDISGRLLFNLMYLAERGVRVR 131
DB 39 GEGCILALDNLDAFAARYLTETWAARTLDVQYIWEEDMSGRLLFSVLSAARGVHR 98

QY 132 LLLDDNNTRGLDLLLLALDHPNIEVRLNFPFLRWKRALGYLTDFFLRNRRMHNKSFTA 191
DB 99 LLLDDNNTPGLDRTLRLDLSHPNIEVRLNFPFRTLRALGYLTDFFLRNRRMHNKSFTA 158

QY 192 DNRATILGRRNIGDEYFKVGEDTVFADLDILATGVSUVGVSHDFDRYWASHAHNATRII 251
DB 159 DGVVTLVGGRRNIGDAYFGAGEEPLFSDLDVMAIGPVVNDVANDFERWRCSSYSTLQQVL 218

QY 252 RSGNIGKGLQAL-----GYNDETSHALLRYRETVEQSPLYOKIOTGRIDWOSVQT 302
DB 219 SLSE-----QELTQRIELPSWNEIDTR-----RYLHKLETQFMADLRGTLPILWAKT 269

QY 303 RLISDTPAKGLDRDRRKPPIAGRLQDALKOPEKSVLVLPYFVPTKSGTDALAKLVQDGI 362

```

```

RESULT 16
QSPGY9_SALPA
ID QSPGY9_SALPA PRELIMINARY; PRT; 495 AA.
AC QSPGY9;
DT 01-FEB-2005 (TRENBLrel. 29, Created)
DT 01-FEB-2005 (TRENBLrel. 29, Last sequence update)
DT 01-FEB-2005 (TRENBLrel. 29, Last annotation update)
DE Hypothetical protein ymdC.
GN Name=ymdC; OrderedLocusNames=SPA1703;
OS Salmonella paratyphi-a.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
OX NCBI_TaxID=54388;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=ATCC 9150;
RX PubMed=15531882; DOI=10.1038/ng1470;
RA McClelland M., Sanderson K.E., Clifton S.W., Latreille P.,
RA Porwollik S., Sabo A., Meyer R., Bieri T., Ozersky P., McLellan M.,
RA Harkins C.R., Wang C., Nguyen C., Berghoff A., Elliott G.,
RA Kohlberg S., Strong C., Du F., Carter J., Kremizki C., Layman D.,
RA Leonard S., Sun H., Fulton L., Nash W., Miner T., Minx P.,
RA Delehaanty K., Fronick C., Magrini V., Nhan M., Warren W., Florea L.,
RA Spieth J., Wilson R.K.;
RT "Comparison of genome degradation in Paratyphi A and Typhi, human-
RT restricted serovars of Salmonella enterica that cause typhoid."
RL Nat. Genet. 36:1268-1274(2004).
DR EMBL; CP000026; AAV77627.1; -; Genomic DNA.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR GO; GO:0008654; P:phospholipid biosynthesis; IEA.
DR InterPro; IPR001736; PLD.
DR Pfam; PF00614; PLDC; 2.
DR SMART; SM00155; PLDC; 2.
DR PROSITE; PS50035; PLD; 2.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 495 AA; 56310 MW; 8BAADE58874983A2 CRC64;

Query Match 40.1%; Score 1089.5; DB 2; Length 495;
Best Local Similarity 47.4%; Pred. No. 1.9e-68;
Matches 221; Conservative 77; Mismatches 147; Indels 21; Gaps 5;

QY 72 GLSDIYLLDDPHEALARAALIESAHSLDQYIWRNDISGRLLFNLMYLAERGVRVR 131
DB 39 GEGCILALDNLDAFAARYLTETWAARTLDVQYIWEEDMSGRLLFSVLSAARGVHR 98

QY 132 LLLDDNNTRGLDLLLLALDHPNIEVRLNFPFLRWKRALGYLTDFFLRNRRMHNKSFTA 191
DB 99 LLLDDNNTPGLDRTLRLDLSHPNIEVRLNFPFRTLRALGYLTDFFLRNRRMHNKSFTA 158

QY 192 DNRATILGRRNIGDEYFKVGEDTVFADLDILATGVSUVGVSHDFDRYWASHAHNATRII 251
DB 159 DGVVTLVGGRRNIGDAYFGAGEEPLFSDLDVMAIGPVVNDVANDFERWRCSSYSTLQQVL 218

QY 252 RSGNIGKGLQAL-----GYNDETSHALLRYRETVEQSPLYOKIOTGRIDWOSVQT 302

```



```

Db 219 SLSE-----QELTORIELPESWYNDEITR-----RYLHKLETSQFMAWLDLDCGTLPLIWAKT 269
Qy 303 RLISDTPAKGLDRDRRKPPIAGRIQDALKQPEKSVYLVSPYFVPTKSGTDALAKLVODGI 362
Db 270 RLLSDPSKSGKAQRHSLPLQRLFDVWMSPTERIDIISAYFVPTRAGVAGQLLNLVRKV 329
Qy 363 DVTVLTNSLQATDVAHVSGVYKRPKLLKAGIKLYELOPNHA--VPATKDKGLTGSSVTS 421
Db 330 KIAITLSLAANDVAVHAGYARWKLLRYGVLEYELKPTREHETAVHORGVLGNSGSS 389
Qy 422 LHAKTFTVDGKRIFIGSNLDPGARLNTENGWVIESPKIAEQMERTLADTPPYAYRV 481
Db 390 LHAKTFTSDGSKVIGSNLDPGARLNTENGWVIESPTLATLHKRFTQSORDAANQLR 449
Qy 482 LDKHNRLQWHD--PATRKYTPNEPEAKLWKRIAAKILSLDPIEGLL 525
Db 450 LDRWGRINWIDROQEEBKVLKKEPATRPFQWRVLRVLAAILPVEWLL 495

RESULT 17
Q6F8A6 ACIAD
ID Q6F8A6 ACIAD PRELIMINARY; PRT; 519 AA.
AC Q6F8A6;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Putative phospholipase D protein.
GN OrderedLocusNames=ACIAD3001;
OS Acinetobacter sp. (strain ADP1).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Moraxellaceae; Acinetobacter.
OX NCBI_TaxID=62977;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=ADP1;
RX PubMed=15514110; DOI=10.1093/nar/gkh910;
RA Barbe V., Valletet D., Fonknechten N., Kreimeyer A., Oztas S.,
Labarre L., Cruveiller S., Robert C., Duprat S., Wincker P.,
Ornston L.N., Weissenbach J., Marliere P., Cohen G.N., Medigue C.;
RT "Unique features revealed by the genome sequence of Acinetobacter sp.
ADP1, a versatile and naturally transformation competent bacterium.";
RL Nucleic Acids Res. 32:5766-5779 (2004).
DR EMBL; CR543861; CAG69709.1; -; Genomic_DNA.
DR GO; GO:0016740; P:transferase activity; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR001736; PLD.
DR Pfam; PF00614; PLDC; 2.
DR SMART; SM00155; PLDC; 2.
DR PROSITE; PS00035; PLD; 2.
KW Complete proteome.
SQ SEQUENCE 519 AA; 59913 MW; C87537724E490292 CRC64;

Query Match 34.6%; Score 940.5; DB 2; Length 519;
Best Local Similarity 38.2%; Pred. No. 7.2e-58;
Matches 207; Conservative 101; Mismatches 177; Indels 57; Gaps 10;

Qy 18 MKTSLSLCLLCSCSWLPPEERTESHFNTSKPVLNDILQIRHTPHNGLSDIY 77
Db 1 MLTSLVTVGVINGCSTLPKHGP-EVTQYVIDIDTSQTSQAIIITPLKQ--QNLGLTGYH 57
Qy 78 LLDDPHEALAAARALIESAHSLSLDLQYIWRNDISGRLLFNLMYLAERGVRVRLLLDDN 137
Db 58 MLYEPLAALAAARLIDIAKELTDLDQYIWRNDIKIGALHALTRAADRGVRLLLDDN 117
Qy 138 NTRGLDDLLALLDGHNPINLEVRFPFLVRKWRALGYLTFDFFLRNRRMHNKSFADNRATI 197
Db 118 NAKQIEGVLLALSQHONIEVKLFNPFYRFRKYRAMDMILDKRINRRMHNKSFADNQVAL 177
Qy 198 LGGRNIGDEYFKVGEDTVPADLLIATGTSVGEVSHDDFDRWASHAHNATRIIRSGNIG 257
Db 178 IGGRNMTNYNNSDNQYFSDVLLVGAADVDDIVKSFDDYWNDEYAYSVQNIIV----- 231
Qy 258 KGLQALGYNDETSRHALRYRETVEQ-SPLYQKIQTG-----RID 296

```

```

Db 232 -----NFEQHR---LRYQSLKQQLDDYQEVTVQVNYLDLTTRSHAFQWLNNNQLD 280
Qy 297 WQSVQTRLIISDTPAKGLDRDRRKPPIAGRIQDALKQPEKSVYLVSPYFVPTKSGTDALAK 356
Db 281 W--VKATVVKDAPDKIRAKANKEQHNLNFQLIKHEKPEDNVDLISAYFVPEKQAKHLS 338
Qy 357 LVQDGIDVTVLTNSLQATDVAHVSGVYKRPKLLKAGIKLYELOP-----NHA 405
Db 339 MAQDGIKRVLTNSFKANDVPVVFHAFYAKYRQELLENGVOLYEFPLTEKEELDXYTDEL 398
Qy 406 VPATKDKGLTGSSVTSLHAKTFTVDGKRIFIGSNLDPGARLNTENGWVIESPKIAEQM 465
Db 399 VKAK-VSIKGLSSSLHAKWALDQKQVIGSFNDFPRSAYLNTLGVILNSPLAMAV 457
Qy 466 ERTLADTPPYAYRVLTLDKHNRLQWHDPA--RTKYPNEPEAKLWKRIAAKILSLPIEG 523
Db 458 HQTMDQQLTHYAYKLVLDANHKINWYRQTSTGTKIYTKPRMKWQKAGIKLISWLPTEG 517
Qy 524 LL 525
Db 518 FM 519

RESULT 18
Q4LR64_9BURK
ID Q4LR64_9BURK PRELIMINARY; PRT; 516 AA.
AC Q4LR64;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Phospholipase D/Transphosphatidylase precursor.
GN ORFNames=Bcen242DRAFT_3286;
OS Burkholderia cenocepacia H12424.
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Burkholderiaceae; Burkholderia; Burkholderia cepacia complex.
OX NCBI_TaxID=331272;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=H12424;
RG US DOE Joint Genome Institute (JGI-PGF);
RA Copeland A., Lucas S., Lapidus A., Barry K., Detter C., Glavina T.,
Hammon N., Israni S., Pitluck S., Richardson P.;
RT "Sequencing of the draft genome assembly of Burkholderia cenocepacia
H12424.";
RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=H12424;
RG US DOE Joint Genome Institute (JGI-ORNL);
RA Larimer F., Land M.;
RT "Annotation of the draft genome assembly of Burkholderia cenocepacia
H12424.";
RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
preliminary data.
DR EMBL; AAHL01000025; EAM18523.1; -; Genomic_DNA.
KW SIGNAL.
FT SIGNAL
SQ SEQUENCE 516 AA; 57599 MW; 56B9D87A8D18A60E CRC64;

Query Match 32.4%; Score 881.5; DB 2; Length 516;
Best Local Similarity 40.5%; Pred. No. 1.1e-53;
Matches 214; Conservative 77; Mismatches 203; Indels 35; Gaps 11;

Qy 18 MKTSLSLCLLCSCSWLPPEERTESHFNTSKPVLNDILQIRHTPHNGLSDIY 77
Db 2 IRLSWGAISLLMLAACASLPPOADRAPTHAFATDTRGLGVAFRQQAHH-PQDAFH 60
Qy 78 LLDDPHEALAAARALIESAHSLSLDLQYIWRNDISGRLLFNLMYLAERGVRVRLLLDDN 137
Db 61 LLTDPVDALDARVLLADRAERSLDLQYIYWHDDLTHGELADAVMRAADRGVRVRLLLDD 120

```



```

Qy 140 RGLDLLLALDSHNIENIENLPPVILRKWRALGYLTDPFLNRRMHNSFTADNRATILG 199
Db 122 RD-DQVWALLNDHNIENIENLPPVILRKWRALGYLTDPFLNRRMHNSFTADNRATILG 180
Qy 200 GRNIGDEYFKVGEDTVFADLDILATGSSVVGSHDFDRYWAHSHAHNATRIIRSGNICKG 259
Db 181 GRNIGNEVFSTQVFGDFDLVLFEGPAVVTATQFDLYNSDQTSVSALVSLSDVETQ 240
Qy 260 LOALGYNDSTSRHALLRYR--TVQSPPLYQKIQTGRIDMQSVQTRLISDTPAKGLDRDR 318
Db 241 YALEDLVDVNALEAFPRDGEYDISQELFEHLKHGTLKLYNGEAVLWYDLDPKVEYTRDSQ 300
Qy 319 KPPTAGRLQDALKOPEKSVILVSPYFPTKSGTALAKLVODGIDVTVLNLSQATDVAA 378
Db 301 ---VAEQALERNVNTSILISPYFVTEAGTGLKRLAESGVKVTIVTNSLASNDVFA 357
Qy 379 VHSQVVKRKLKAGIKLYLQPNHAPVATKDKGLTSSVTSIHAFTFIVDGKRIFGS 438
Db 358 VHGYAKYRELLSGITLWIKASAEI--KRWSLIGSQRLASHAKVIVIDDADFVGS 415
Qy 439 FNLDPRLNTGEMGVVIESPKIAEQMERTIADTTPPEYAYRVTLDKHNRLQWHPATRKT 498
Db 416 MNWDPERSAYINTEMAVHITHEPYVCKATVQLNSALPKDAYVVEIREGN-LHMTDLGNKV 474
Qy 499 YPNPEAKLWKRIAKILSLPIEGLL 525
Db 475 YDSEPFESSIWRSIGAWLSGVLPFIEGML 501

```

RESULT 21

```

Q8D861.VIBVU
ID Q8D861.VIBVU PRELIMINARY; PRT; 524 AA.
AC Q8D861;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Phospholipase D-family protein.
GN OrderedLocusNames=VVI3126;
OS Vibrio vulnificus.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Vibrio.
OX NCBI_TaxID=672;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=CMCP6;
RA Rhee J.H., Kim S.Y., Chung S.S., Kim J.J., Moon Y.H., Jeong H.,
  Choy H.E.;
RT "Complete genome sequence of Vibrio vulnificus CMCP6.";
RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE016807; AA011447.1; -; Genomic DNA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR001736; PLD.
DR Pfam; PF00614; PLDC; 2.
DR SMART; SM00155; PLDC; 2.
DR PROSITE; PS00035; PLD; 2.
KW Complete proteome.
SQ SEQUENCE 524 AA; 59929 MW; 8C88CBC38C8CADA7 CRC64;

```

```

Query Match 32.0%; Score 871.5; DB 2; Length 524;
Best Local Similarity 39.4%; Pred. No. 5.6e-53;
Matches 211; Conservative 89; Mismatches 197; Indels 39; Gaps 14;

Qy 5 PKIQAMPSETISPMKT---RSLISLLCLLLCSCSWLPPLEERTESRHFNTPSKPVL--- 58
Db 7 PEYLNKPPFGLGPLRTFIHKALTYCSVLLCACT---LENHSP---FDKQPSVOLGYQ 59

Qy 59 -DNILQ--IRUTPHN-NGLSDIYLLDDPHEALARAALIESAHSLSLDIQYIWRNDISGR 114
Db 60 ADSRLSAYLNHPQDRENLTAFPLDKGHDALLARLALIEAADKTLDIQYIYFRNDETQ 119

Qy 115 LLFNLMLYLAERGVRVRLDDNNTRGUDLLALLDSDHPNIEVRLFNPFVLRKWRALGYL 174

```

```

Db 120 LLTWRLFEARQGRVVRILLDDDMQKRN-DEGLARLSAHPNIQIRLFNPHQVKTARTLAMA 178
Qy 175 TDFPRLNRMHNSFTADNRATILGGRNIGDEYFKVGEDTVFADLDILATGSSVVGVEVSHD 234
Db 179 SDFSLNRMHNSLTADSVVAIVGGRNIGNEVFSVNSPVEFGDFDLMLYGNSEVQTAEQ 238
Qy 235 FDRYWAHSHA-----HNATRIIRSGNIGKGLQALGYNDSTSRHALLRYRVEVQSPL 286
Db 239 FDLYNSLHVAPIEWLTDNPIPVTE----ELQA--WLKETQLEQKFTQGRYDFTQLQL 291
Qy 287 YQKIQTGRIDMQSVQTRLISDTPAKGLDRDRRKPIAGRLQDALKOPEKSVILVSPYFVP 346
Db 292 YQQTDFKSLVYWGKGQWYDLDPK---VDTQAPQADNLASLLRTVKDSLVLISPYFVP 348
Qy 347 TKGSTDALAKLVODGIDVTVLNLSQATDVAAVHSGYVYKRPKLLKAGIKLYELQPNHAV 406
Db 349 TERGTQALVEAHRGVDTITVTNSLAGNDVFAVHWYAKYRQDLVEAGIQLWETKASARI 408
Qy 407 PATKDKGLTSSVTSIHAFTFIVDGKRIFGSNLDPRLNTGEMGVVIESPKIAEQME 466
Db 409 DS--KWSFTGSSRSSLSHAKVLLIDHRLLFAGSMNWDPRSLALNTTEMAAVIEHPDYVOSSE 466
Qy 467 RTLADTTPEYAYRVTLDKHNRLQWHPATRKYTPNEPEAKLWKRIAKILSLPIE 522
Db 467 AKLPMGLETNAYQVRM-KNGEVAWFDHQSWFDEPEATVWRKIGAWFAFILPIE 521

```

RESULT 22

```

Q7MMB3.VIBVY
ID Q7MMB3.VIBVY PRELIMINARY; PRT; 554 AA.
AC Q7MMB3;
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein VVI1159.
GN OrderedLocusNames=VVI1159;
OS Vibrio vulnificus (strain YJ016).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Vibrio.
OX NCBI_TaxID=196600;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX PubMed=14656965; DOI=10.1101/gr.1295503;
RA Chen C.-Y., Wu K.-M., Chang Y.-C., Chang C.-H., Tsai H.-C.,
  Liao T.-L., Liu Y.-M., Chen H.-J., Shen A.B.-T., Li J.-C., Su T.-L.,
  Shao C.-P., Lee C.-T., Hor L.-I., Tsai S.-F.;
RT "Comparative genome analysis of Vibrio vulnificus, a marine
  pathogen.";
RL Genome Res. 13:2577-2587(2003).
DR EMBL; BA000037; BAC93923.1; -; Genomic DNA.
DR GO; GO:0003824; F:catalytic activity; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR001736; PLD.
DR Pfam; PF00614; PLDC; 2.
DR PROSITE; PS00035; PLD; 2.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 554 AA; 63341 MW; 8A7980BB096AB6C8 CRC64;

```

```

Query Match 32.0%; Score 869.5; DB 2; Length 554;
Best Local Similarity 38.2%; Pred. No. 8.4e-53;
Matches 207; Conservative 90; Mismatches 194; Indels 51; Gaps 12;

Qy 5 PKIQAMPSETISPMKT---RSLISLLCLLLCSCSW--LPPLEBTE-----SR 48
Db 37 PKYLNKPPFGLGPLRTFIHKALTYCSVLLCACTSFENHSFPDKQPSYQLGYQADSRLSA 96

Qy 49 HFNTSKPVLDDNLIQIRHTPHNGLSDIYLLDDPHEALARAALIESAHSLSLDIQYIWR 108
Db 97 YLN-----HHQDRENLTAFPLDKGHDALLARLALIEAADKTLDIQYIYFR 143

Qy 109 NDISGRLLFNMLYLAERGVRVRLDDNNTRGUDLLALLDSDHPNIEVRLFNPFVLRKW 168

```

Db 144 NDETQQLTWLFEAAQGVVRVRIILLDDMQKN-DEGLARLSAHFNIQIRLFENPHQYRTA 202
Qy 169 RALGYLTDPFRLNRRMHNKSFADNRATILGRNIGDVEYFKVGEDTVFADLDILATGSVV 228
Db 203 RFLWASDFSLNRRMHNKSLTADSVVAIVGRNIGNEYFVNSPVEFGDFDLMDYNSV 262
Qy 229 GEVSHDFRYWASHA-----HNATRIIRSGNIGKGLQALGYNDETS-RHALLRYRET 280
Db 263 EQTAEQFLYNSLHVAPIEWLTNPVTE-----BELQA--WLKETOLEKQKTOGRYD 315
Qy 281 VEOSPLYOKIOTGRIDWQSVOTRLISDPKALDRDRKPPAIGRLQDALKQPEKSVLV 340
Db 316 FTQLQLYQFTDKSLVFWGQVWYDLPDK---VDTQEQQLADNLTSLLTFRVDSLVLI 372
Qy 341 SPYFVPTKSGDALAKLVQDGDIVTLNLSQATDVAHVHSGYVKYRKLKAGIKLYEL 400
Db 373 SPYFVPTERGTOALVEAQRGVEITIVTNSLASNDVFAVHGWAYKYRQDLVEAGIQWET 432
Qy 401 QPNHAPATKDKGLTGSVTSLSHAKTFIVDGKRFIFGSFNLDPDRSARLNTMGVVIISP 460
Db 433 KASARIDS--KWSFTGSSRSLSHAKVLLIDHRLLFAGSMWDPDRSALLNTEMAIVIEHPD 490
Qy 461 IAEQWERTLATTEYAYRVTLDKHRLQWHDPAKTRKTYPNPEAKLWKRIAAKILSLP 520
Db 491 YVQSSEAKLPMGLTNAVQVRM-KNGEVANFQHSQVWFDSEPEATVWRKIGANFAGILP 549
Qy 521 IE 522
Db 550 IE 551

RESULT 23

Q8D800 VIBVU PRELIMINARY; PRT; 501 AA.
ID Q8D800 VIBVU PRELIMINARY; PRT; 501 AA.
AC Q8D800;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DE Phospholipase D-family protein.
GN OrderedLocusNames=VV13197;
OS Vibrio vulnificus.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Vibrio.
OX NCBI_TaxID=672;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=CMCP6;
RA Rhee J.H., Kim S.Y., Chung S.S., Kim J.J., Moon Y.H., Jeong H.,
RA Choy H.E.;
RT "Complete genome sequence of Vibrio vulnificus CMCP6.";
RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: AE016807; AA011511.1; -; Genomic DNA.
DR GO: GO:0016740; P:transferase activity; IEA.
DR GO: GO:0008152; P:metabolism; IEA.
DR InterPro: IPR001736; PLD.
DR Pfam: PF00614; PLDC; 2.
DR SMART: SM00155; PLDC; 2.
DR PROSITE: PS50035; PLD; 2.
KW Complete proteome.
SQ SEQUENCE 501 AA; 56571 MW; BF094CE115F8D30B CRC64;

Query Match 31.7%; Score 862.5; DB 2; Length 501;
Best Local Similarity 38.0%; Pred. No. 2.3e-52;
Matches 197; Conservative 91; Mismatches 198; Indels 33; Gaps 8;
Qy 21 RSLISLLCLLSCSSWLPLEERTESRHNFTSK-PVLLDNILQIRHTPHNGLSDIYLL 79
Db 2 RVLVAVSCFGLACASNPVTEKKLSYHSGYQEGSTLAKYEQLRTPETVKDNLNLTGPHSL 61
Qy 80 DDPHEALAAARALTESAHSLLQYIWRNDISGRLLFNLMYLAERGVRVRLLLDDNNT 139
Db 62 NNGDDALLALIALITSQAHSLLDQYLYFGDDETSKIITWRLYEAAQGVKVRLLDDMQR 121

Qy 140 RGLDLLLLSHDNIEVRLFPNPFVLRKWRALGYLTDPFRLNRRMHNKSFADNRATILG 199
Db 122 RD-DQWALLNDHNIEIRLFPNPFVLRKWRALGYLTDPFRLNRRMHNKSFADNRATILG 180
Qy 200 GRNIGDEYFKVGEDTVFADLDILATGSVGVSHDFRYWASHAHNATRIIRSGNIGK 259
Db 181 GRNIGNEYFVSTQVEFGDFVLLFGPAVVEATQFQDLYWNSDQTPVSAV----- 232
Qy 260 LQALGYNDESRHAL-----LRYRE---TVEQSPLYOKIOTGRIDWQSVOTRLIS 306
Db 233 ----SLSDYQYALEDLVDNALVAFRDEGYDISQLELFEHLKGTGLKYWEAVLWY 288
Qy 307 DTPAKGLDRDRKPPAIGRLQDALKQPEKSVLVSPYFVPTKSGTDLAKLVQGDIVTV 366
Db 289 DLPDKVETRDSQ---VAEQLAERLNNVTDYILIIISPYFVPTKSGTDLAKLVQGDIVTV 345
Qy 367 LTNSLQATDVAHVHSGYVKYRKLKAGIKLYELQPNHAPATKDKGLTGSVTSLSHAKT 426
Db 346 VTNSLASNDVFAVHGWAYKYRQDLVEAGIQWET 403
Qy 427 FIVDGKRFIFGSFNLDPDRSARLNTMGVVIISPKEIAEQWERTLATTEYAYRVTLDKH 486
Db 404 IVIDDADFVGSNMWDPDRSARLNTMGVVIISPKEIAEQWERTLATTEYAYRVTLDKH 463
Qy 487 RLQWHDPAKTRKTYPNPEAKLWKRIAAKILSLPIEGLL 525
Db 464 -LHWTDLGNCKYVDSEPESSIRWSIGAWLSGVLPDIEGML 501

RESULT 24

Q8PIM4 XANAC PRELIMINARY; PRT; 520 AA.
ID Q8PIM4 XANAC PRELIMINARY; PRT; 520 AA.
AC Q8PIM4;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DE 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Cardiolipin synthase.
GN Name=cls;
OS Xanthomonas axonopodis (pv. citri).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC Xanthomonadaceae; Xanthomonas.
OX NCBI_TaxID=92829;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=306 / ATCC 13902 / XV 101;
RX MEDLINE=22022145; PubMed=12024217; DOI=10.1038/417459a;
RA da Silva A.C.R., Ferro J.A., Reinach F.C., Farah C.S., Furlan L.R.,
RA Quaggio R.B., Monteiro-Vitorello C.B., Van Sluys M.A.,
RA Almeida N.F. Jr., Alves L.M.C., do Amaral A.M., Bertolini M.C.,
RA Camargo L.E.A., Camarotte G., Cannavan F., Cardoso J., Chambergo F.,
RA Ciapina L.P., Cicarelli R.M.B., Coutinho L.L., Cursino-Santos J.R.,
RA El-Dorriy H., Faria J.B., Ferreira A.J.S., Ferreira R.C.C.,
RA Ferro M.I.T., Formighieri E.F., Franco M.C., Greggio C.C., Gruber A.,
RA Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,
RA Locali E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,
RA Martins E.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,
RA Moreira L.W., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,
RA Pereira H.A., Rossi A., Sena J.A.D., Silva C., de Souza R.F.,
RA Spinola L.A.F., Takita M.A., Tamura R.B., Teixeira E.C., Tezza R.I.D.,
RA Trindade dos Santos M., Truffi D., Tsai S.M., White F.F.,
RA Setubal J.C., Kitajima J.P.;
RT "Comparison of the genomes of two Xanthomonas pathogens with differing
RT host specificities.";
RL Nature 417:459-463(2002).
DR EMBL: AE011929; AA037716.1; -; Genomic DNA.
DR GO: GO:0003824; F:catalytic activity; IEA.
DR GO: GO:0008152; P:metabolism; IEA.
DR InterPro: IPR001736; PLD.
DR Pfam: PF00614; PLDC; 2.
DR SMART: SM00155; PLDC; 2.
DR PROSITE: PS50035; PLD; 2.
KW Complete proteome.
SQ SEQUENCE 520 AA; 57436 MW; 61A2DE5E0349A2FA CRC64;

Query Match	31.3%;	Score 850.5;	DB 2;	Length 520;
Best Local Similarity	37.9%;	Pred. No. 1.7e-51;		
Matches 206;	Conservative 80;	Mismatches 182;	Indels 75;	Gaps 15;
y	24	ISLLCLLCS-	-----SSMLPPEERTSRHFNTSPVLLDNI--LQ 63	
b	12	IALLVIVLAS	LSLVYGRFADQRGPVSHVLP------TAVATPIDKVVAPLQ 60	
y	64	IRHTPHNGLSDI	YLLDDPHEALAAARAALIESAEHSLDQYIWRNDISGRLLFNMYLA 123	
b	61	QAHADQ----	TGMVILPDNVDAFAVRAUTARAAGSLDQYIWHADFTGNLLHNELLRA 116	
y	124	AERGVRVRLLLD	DNTRGLDLLLLALDHPNIEVRLFPVLRK---WRALGYLTDPRRL 180	
b	117	ADRGVRVRLLLD	DMNIHGSDSVLAALDHPLEIETIRLFPNTRAREGTLMRGVMVLRMFSI 176	
y	181	NRMRNKSPFADNR	ATILGGRNIGDEYFKVGEDTVFADLDILATGSVVVGEVSHDFDRYWA 240	
b	177	NRMRNKAWIDGR	TAVVGGRVGDEYFDAARDTNFMDMDAALMGPAVGQAEQVFDAYWN 236	
y	241	SHSANNATRIIRS	-----GNTKGQLQALGYNDTSRHALLRYRETVEQSPYOKIQI 291	
b	237	SPNALPLSLVAK	PQALEALRGLSDAGWASVRAHP-----YVERLKRSPSVRELM 287	
y	292	TG--RIDMQSVOT	RLISDTPAKG----LDRDRKPIIAGRLQDALKQPEKSVLWVSPYFV 345	
b	288	QDRPVHWS--QAR	IVSDPPEKAGAPQADMTFVLIGEWAHA----QRELKVISPYFV 342	
y	346	PTKSGTDALAKI	VDGIDVTVLITNSLQATDVAAVHSGYVKYRKPILLKAGIKLYELQPNHA 405	
b	343	PGDEGMRWIGH	LQRQNRVRSILTNSLAANDVVAVHSGYAGYRVPLLQLGVRLLHELKP-- 399	
y	406	VPATKDKGLTG	SSVTSLHAKTFIVDGKIFIGSFNLDPRSARLNTMGVWIESPKIAEQM 465	
b	400	-MGKPDGSLFG	SSGASLTKAEFVDDSGSGFTGSFNLDPRSMLNTEMLGLLDFNDRVVTAE 458	
y	466	ERTL-ADTTPYAY	RVRTLDKHNRLQWHPDAPTR--KTYPNEPEAKLWKRIAAKILSLPLIE 522	
b	459	ERLYNHKVSAP	YSYRVTLDR-CELWHDDAAQPEVWSREPAASVWRGGAAMVWGLPLE 517	
y	523	GLL 525		
b	518	SQL 520		
RESULT 25				
6MQL7_BDEBA PRELIMINARY; PRT; 530 AA.				
C	C	Q6MQL7_BDEBA PRELIMINARY; PRT; 530 AA.		
C	C	Q6MQL7;		
D	D	Q6MQL7;		
T	T	05-JUL-2004 (TrEMBLrel. 27, Created)		
T	T	05-JUL-2004 (TrEMBLrel. 27, Last sequence update)		
T	T	05-JUL-2004 (TrEMBLrel. 27, Last annotation update)		
E	E	Putative phospholipase D precursor.		
N	N	OrderedLocustNames=Bd0448;		
S	S	Bdellovibrio bacteriovorus.		
C	C	Bacteria; Proteobacteria; Deltaproteobacteria; Bdellovibrionales;		
C	C	Bdellovibrionaceae; Bdellovibrio.		
X	X	NCBI_TaxID=959;		
N	N	[1]		
P	P	NUCLEOTIDE SEQUENCE.		
C	C	STRAIN=HD100 / DSM 50701 / ATCC 15356 / NCIB 9529;		
X	X	PubMed=14752164; DOI=10.1126/science.1093027;		
R	R	Rendulic S., Jagtap P., Rosinus A., Eppinger M., Baar C., Lanz C.,		
R	R	Keller H., Lambert C., Evans K.J., Goesmann A., Meyer F.,		
A	A	Sockett R.E., Schuster S.C.		
T	T	"A predator unmasked: life cycle of Bdellovibrio bacteriovorus from a		
T	T	genomic perspective.";		
L	L	Science 303:689-692(2004).		
R	R	EMBL; BX842647; CAE78430.1; -; Genomic DNA.		
R	R	GO; GO:0003824; F:catalytic activity; IEA.		
R	R	GO; GO:0008152; P:metabolism; IEA.		
R	R	InterPro: IPR001736; PLD.		

[illegible]

RESULT 25

C	Q6MQL7_BDEBA	PRELIMINARY;	PTT;	530 AA.
C	Q6MQL7;			
T	C	05-JUL-2004	(trEMBLrel. 27, Created)	
T	T	05-JUL-2004	(trEMBLrel. 27, Last sequence update)	
T	T	05-JUL-2004	(trEMBLrel. 27, Last annotation update)	
E	E	Putative phospholipase D precursor.		
S	S	OrderedLocusNames=Bd0448;		
S	S	Bdellovibrio bacteriovorus.		
C	C	Bacteria: Proteobacteria; Deltaproteobacteria; Bdellovibrionales;		
C	C	Bdellovibrionaceae; Bdellovibrio.		
X	X	NCBI_TaxID=959;		
N	N	[1]		
P	P	NUCLEOTIDE SEQUENCE.		
C	C	STRAIN=HD100 / DSM 50701 / ATCC 15356 / NCIB 9529;		
X	X	PubMed=14752164; DOI=10.1126/science.1093027;		
A	A	Rendulic S., Jagtap P., Rosinus A., Eppinger M., Baar C., Lanz C.,		
A	A	Keller H., Lambert C., Evans K.J., Goesmann A., Meyer F.,		
A	A	Sockett R.E., Schuster S.C.		
T	T	"A predator unmasked: life cycle of Bdellovibrio bacteriovorus from a		
T	T	genomic perspective."		
L	L	Science 303:689-692(2004).		
R	R	EMBL; BX842647; CAE78430.1; -; Genomic DNA.		
R	R	GO; GO:0003824; F:catalytic activity; IEA.		
R	R	GO; GO:0008152; P:metabolism; IEA.		
R	R	InterPro: IPR001736; PLD.		

```

RESULT 26
Q4FTL4_9GAMM
ID Q4FTL4_9GAMM PRELIMINARY; PRT; 570 AA.
AC AC
Q4FTL4;
DT 13-SEP-2005 (T-EMBLrel. 31, Created)
DT 13-SEP-2005 (T-EMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (T-EMBLrel. 31, Last annotation update)
DE Probable phospholipase D.
GN ORFNames=Psyc 0791;
OS Psychrobacter_arcticum 273-4.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Moraxellaceae; Psychrobacter.
OX NCBI_TaxID=259536;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=273-4;
RG NASA Astrobiology Institute;
RA Ayala-del-Rio H.L., Chain P., Ponder M.A., Di Bartolo G., Ivano
RA Bergholz P.W., Hauser L., Land M., Bakermans C., Rodrigues D.,
RA Klappenschach J.A., Zarka D., Larimer F., Richardson P., Thomas
RA Tiedje J.M.;
RT "Complete sequence of Psychrobacter arcticum 273-4.";
RL Submitted (JUL-2005) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.

```



```
QY 62 LQIRHTPHNGLSDIYLLDDPHEALAAARALIESAEHSLDQYYIWRNDISGRLLFNLMY 121
Db 59 LQQAHA-----GKTGMVILSDNIDAFVAVRALTARAAGRSLDQYYIWHADFNGLLHNE 114
QY 122 LAAERGVVRLLDDNNTRGLDDLLALDHPNIEVLPNPFVLRK---WRALGYLTDFF 178
Db 115 RAADRGVVRLLDDNNHGHSDSVLAALDHPLEIRLFPNTRAREGTLMRGVEMVLRMF 174
QY 179 RLNRMHKNSFTADNRATILGGRNIGDEYFKVGDFTVFADLDILATGSGVGEVSHDFDY 238
Db 175 SINRRMKNKAWIADGRIAVVGRNVGDEYFDAARDTNFMDMDAALIGPTVQSEQVFDAY 234
QY 239 WASHSAHNATRIIRS-----GNIGKGLQALGYNDETSRHALLRYRETVEQSPLYQK 289
Db 235 WNSPNALPLAALVTAKPOALEQLRSLDAGMASVGAHP-----YVERLRQSPSVHA 285
QY 290 IQTGR--IDWQSVQTRLLISDTPAKG---LDRDRKPPIAGRLQDALQKPEKSVYLVSPY 343
Db 286 LMQGNRPVHWVQ-QARIVADPPEKAEGAPPDWMPTILIGEMANA-----RRELKLI 340
QY 344 FVPTKSGTDALAKLVQDGDIVTLTNSLQATDVAHVSGYVKYRKPLKAGIKLYELQPN 403
Db 341 FVPGDDGMRWIGGLRKDRVSVLTNSLAANDVVAHVSGYADYRIPLLQQGVRLHELKP- 399
QY 404 HAVPATKDKGLTGSVTSLSHAKTFIVDGKRIFIGSFNLDPRSARLNTMGVVISP KIAE 463
Db 400 ---MGRPDSGLFGSSGASLHTKAFVVDSDSAGFIGSFNLDPRAMNLTMGLLFEDRTVTA 456
QY 464 QMERTL-ADTTPYAYRVTLDKHNLQWHDPATR--KTPYNEPEAKLWKRTAAKILSLP 520
Db 457 ELERLYNHKVSAPVSYRVTLQO-GGLRWHDDAAQPPAVWTREPAASVWRRGAAGVGLWLP 515
QY 521 IEGLL 525
Db 516 LDSQL 520

RESULT 29
Q8P7A5_XANCP
ID Q8P7A5_XANCP PRELIMINARY; PRT; 520 AA.
AC Q8P7A5
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-WAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Cardiolipin synthase.
GN Name=cls;
OS Xanthomonas campestris (pv. campestris).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC Xanthomonadaceae; Xanthomonas.
OX NCBI_TaxID=340;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=ATCC 33913 / NCPPB 528;
RX MEDLINE=2022145; PubMed=1024217; DOI=10.1038/417459a;
RA da Silva A.C.R., Ferro J.A., Reinach F.C., Farah C.S., Furlan L.R.,
RA Quaggio R.B., Monteiro-Vitorello C.B., Van Sluys M.A.,
RA Almeida N.F. Jr., Alves L.M.C., do Amaral A.M., Bertolini M.C.,
RA Camargo L.E.A., Camarotte G., Cannavan F., Cardozo J., Chambergo F.,
RA Ciapina L.P., Cicarelli R.M.B., Coutinho L.L., Cursino-Santos J.R.,
RA El-Dorzy H., Faria J.B., Ferreira A.J.S., Ferreira R.C.C.,
RA Ferro M.I.T., Formighieri E.F., Franco M.C., Greggio C.C., Gruber A.,
RA Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,
RA Locali E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,
RA Martins E.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,
RA Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,
RA Pereira H.A., Rossi A., Sena J.A.D., Silva C., de Souza R.F.,
RA Spinola L.A.F., Takita M.A., Tamura R.E., Teixeira E.C., Tezza R.I.D.,
RA Trindade dos Santos M., Truffi D., Tsai S.M., White F.F.,
RA Subetal J.C., Kitajima J.P.;
RT "Comparison of the genomes of two Xanthomonas pathogens with differing
RT host specificities."
RL Nature 417:459-463(2002).
```

```
DR EMBL; AE012383; AAM41978.1; -, Genomic DNA.
DR GO; GO:0003824; F:catalytic activity; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR001736; PLD.
DR Pfam; PF00614; PLDC; 2.
DR SMART; SM00155; PLDC; 2.
DR PROSITE; PS00035; PLD; 2.
DR Complete proteome.
SQ SEQUENCE 520 AA; 57165 MW; 8849CFED240410A3 CRC64;

Query Match 31.0%; Score 843; DB 2; Length 520;
Best Local Similarity 40.4%; Pred. No. 5.8e-51;
Matches 196; Conservative 71; Mismatches 174; Indels 44; Gaps 12;

QY 62 LQIRHTPHNGLSDIYLLDDPHEALAAARALIESAEHSLDQYYIWRNDISGRLLFNLMY 121
Db 59 LQQAHA-----GKTGMVILSDNIDAFVAVRALTARAAGRSLDQYYIWHADFNGLLHNE 114
QY 122 LAAERGVVRLLDDNNTRGLDDLLALDHPNIEVLPNPFVLRK---WRALGYLTDFF 178
Db 115 RAADRGVVRLLDDNNHGHSDSVLAALDHPLEIRLFPNTRAREGTLMRGVEMVLRMF 174
QY 179 RLNRMHKNSFTADNRATILGGRNIGDEYFKVGDFTVFADLDILATGSGVGEVSHDFDY 238
Db 175 SINRRMKNKAWIADGRIAVVGRNVGDEYFDAARDTNFMDMDAALIGPTVQSEQVFDAY 234
QY 239 WASHSAHNATRIIRS-----GNIGKGLQALGYNDETSRHALLRYRETVEQSPLYQK 289
Db 235 WNSPNALPLAALVTAKPOALEQLRSLDAGMASVGAHP-----YVERLRQSPSVHA 285
QY 290 IQTGR--IDWQSVQTRLLISDTPAKG---LDRDRKPPIAGRLQDALQKPEKSVYLVSPY 343
Db 286 LMQGNRPVHWVQ-QARIVADPPEKAEGAPPDWMPTILIGEMANA-----RRELKLI 340
QY 344 FVPTKSGTDALAKLVQDGDIVTLTNSLQATDVAHVSGYVKYRKPLKAGIKLYELQPN 403
Db 341 FVPGDDGMRWIGGLRKDRVSVLTNSLAANDVVAHVSGYADYRIPLLQQGVRLHELKP- 399
QY 404 HAVPATKDKGLTGSVTSLSHAKTFIVDGKRIFIGSFNLDPRSARLNTMGVVISP KIAE 463
Db 400 ---MGRPDSGLFGSSGASLHTKAFVVDSDSAGFIGSFNLDPRAMNLTMGLLFEDRTVTA 456
QY 464 QMERTL-ADTTPYAYRVTLDKHNLQWHDPATR--KTPYNEPEAKLWKRTAAKILSLP 520
Db 457 ELERLYNHKVSAPVSYRVTLQO-GGLRWHDDAAQPPAVWTREPAASVWRRGAAGVGLWLP 515
QY 521 IEGLL 525
Db 516 LDSQL 520

RESULT 30
Q6G185_BAROU
ID Q6G185_BAROU PRELIMINARY; PRT; 523 AA.
AC Q6G185;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Phospholipase D-family protein.
GN OrderedLocusNames=B001680;
OS Bartonella quintana (Rochalimaea quintana).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Bartonellaceae; Bartonella.
OX NCBI_TaxID=803;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=Toulouse;
RX PubMed=15210978; DOI=10.1073/pnas.0305659101;
RA Almark U.C.M., Frank A.C., Karlberg E.O., Legault B.-A., Ardell D.H.,
RA Canbaeck B., Eriksson A.-S., Naeslund A.K., Handley S.A., Huvet M.,
RA La Scola B., Holmberg M., Andersson S.G.E.;
RT "The louse-borne human pathogen Bartonella quintana is a genomic
RT derivative of the zoonotic agent Bartonella henselae.";
```


RL	Proc. Natl. Acad. Sci. U.S.A.	101:9716-9721 (2004).
DR	EMBL; BX897700; CAF25671.1; -; Genomic DNA.	
DR	GO: GO:0003824; F:catalytic activity; IEA.	
DR	GO: GO:0008152; P:metabolism; IEA.	
DR	InterPro: IPR001736; PLD.	
DR	Pfam; PF00614; PLDC; 2.	
DR	SMART; SM00155; PLDC; 2.	
DR	PROSITE; PS0035; PLD; 2.	
KW	Complete proteome.	
SO	SEQUENCE 523 AA; 59265 MW; BDCE70FC51BDC16 C	

Query Match	30.4%;	Score 828;	DB 2;	Length 523;
Best Local Similarity	43.2%;	Pred. No. 6.7e-50;		
Matches	200;	Conservative 64;	Mismatches 161;	Indels 38; Gaps 13;
Qy	84	EALAAARALIESAEHSLDQYIWRNDISGRLLFNMLYLAERGVRVRLLLDDNTRGLD	143	
Db	78	DAFCVRAIGAAQAGRSJOLMYIWNDDLTGRLLSIEVEAADRGVRVRLLLDDINAQARD	137	
Qy	144	DLIALDLSHNIENVRLENPFVLRKW---RALGYLTFPRLNRRMHNKSFTADNRATILGG	200	
Db	138	PAYALDKHPRIEVRMFPNPGSRKGLRGLEILRAITVTRRMHNKAFIVDGRVAFVGG	197	
Qy	201	RNIGDEYFKVGEDTVFADLDILATGSSVVGVEHSHFDRIYWAHSAHNATRIIRSGNIGKL	260	
Db	198	RNLADSIFYDAGEESHFRDLDMLIGPSVKVYETIFFDDFNWSAVVLPIHTLV----VPKST	253	
Qy	261	QALGYNDETSRHALLRRET--VEOSPL-YQK-----IQTRIDWQSVQTRLISITPA	310	
Db	254	SDLGYMKEKLUR---KFRDSKVAKVLYDYVKEHINFDCFIQTGRKRLFLADKVVVLSDPPE	309	
Qy	311	KGILDRDRKPIAGR-LQDALKQ---PEKSVLYSVFYFVPTKSGTDALAKLVQDGDIVT	365	
Db	310	KAL---RKK---AGNLWMKALSQVIGDAQKTQVITSPYFVGKVTGQSFRNLVSKGVDVK	363	
Qy	366	VLNLSLQATDVAAVHSGVVKYRKPLLKAGIKLYELQPNHAPVATKDGLTGSSVTSLHAK	425	
Db	364	ILTNLSLAATDVALVHGQVYPRKALLKSGVKLYELKPDGNVHGLR---LFRSSKAILHTK	420	
Qy	426	TFIVDGKRIFIGSNLDRSARLNTMGWVIESPKIAEQMERTUA-DITTPYAYRVTLDK	484	
Db	421	AFLVDRKTAFIGSLNFDPSAALNTMGWILFECAPITARLDLSPSEETGEMSYNLRDNL	480	
Qy	485	HNRLQWHDPAIRTKTY---PNPEAKLVKRIAAKILSLPIEGLL	525	
Db	481	NNRIYDFIENEKQYSIDYSPESNFWRRATRIISWLPISQOL	523	

```

RESULT 31
ID Q6G4X0 BARHE
IQ Q6G4X0 BARHE PRELIMINARY; PRT; 523 AA.
DC Q6G4X0;
AC 05-JUL-2004 (TReMBLrel. 27, Created)
DT 05-JUL-2004 (TReMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TReMBLrel. 27, Last annotation update)
DE Phospholipase d-family protein.
OS OrderedLocusNames=BH01790;
GN Bartonella henselae (Rochalimaea henselae).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Bartonellaceae; Bartonella.
OX NCBI_TaxID=38323;
[1]
NU NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RP STRAIN=ATCC 49882 / Houston 1;
RC PubMed=15210378, DOI=10.1073/pnas.0305659101;
RX Alsmark U.C.M., Frank A.C., Karlberg E.O., Legault B.-A., A
RA Canbaeck B., Eriksson A.-S., Naeslund A.K., Handley S.A., H
RA La Scala B., Holmberg M., Andersson S.G.E.;
RT "The louse-borne human pathogen Bartonella quintana is a ge
RT derivative of the zoonotic agent Bartonella henselae."
RL Proc. Natl. Acad. Sci. U.S.A. 101:9716-9721 (2004) .
DR ENBL; BX897659; CAF26991.1; -; Genomic DNA.
DR GO; GO:0003824; F:catalytic activity; IEA.

```

DR	GO; GO:0008152; P:metabolism; IEA.
DR	InterPro; IPR001736; PLD.
DR	Pfam; PF00614; PLDC; 2.
DR	SMART; SM00155; PLDC; 2.
DR	PROSITE; PS00035; PLD; 2.
KW	Complete proteome.
SQ	SEQUENCE 523 AA; 59379 MW; 262FCDCC8C269D5AD CRC64;
	Query Match 30.3%; Score 825; DB 2; Length 523;
	Best Local Similarity 41.6%; Pred. No.1.le-49;
	Matches 191; Conservative 73; Mismatches 165; Indels 30; Gaps
Qy	84 EALAARAALTESAEHSIDLOYIWRNDISGRLLFNMLYLAERGVVRVRLLLDDNTRGLD 143
Dd	78 DAFCVRAAGAAKGRSLDMYYYIWDDDLTGRLLSSEIVEAADRGVRVRLDDINAQARD 137
Qy	144 DLLLALDSHPNIIEVLRFNFVLRKW---RALGYLTDFPLNRMRHMKSFADNRNATILGG 200
Dd	138 PAYMALDKPHIEVRMFENPGSRKGVRGLEIIIRAITVTVRMHKAFIVDGRMAFVG 197
Qy	201 RNICDEYFKVGEDVFADLDILATGSVVGEVSHPDFRYWAHSAHNATRIRSGNIGKL 260
Dd	198 RNLASDYFDAGKESHFRDLMLICGPSVKVENIFDDFWNSAVVLPPIHTLV----VPKSV 253
Qy	261 QALGCYNDETSRHALLRYRETVEQSPLYQK-----TOTGRIHQWSQVTRLISDTP 309
Dd	254 NDIAF----WREKURQFRDS-KAKAYLDYVKKHISPDCTQTGKRFLADKVVSDDPP 308
Qy	310 AKGLDRRRKPPIAGRILDALKQPEKSYYLVSPFYVPTKSTGDALAKLVQGDIVTLTN 369
Dd	309 EKVL-RKKAENWLKALSQVIDAQKTVOITSPYFVGKVGTONFSNLVSKGVDVKILT 367
Qy	370 SLQNTDVAAHVHSGVVKRKPLLKAGIKLYELOPHNAVDPATKDGLTGS SVTSLLHAKTIV 429
Dd	368 SLAATDVALVHGCVYPYRKMLLKSGVKLYELPGGNMHLR--LFRSSKASLHTKAFIV 424
Qy	430 DGKRIFGTSFNLDRPSARLNTMGWIESPKIAQMERTLA-DTTPYAYRVRLTDKNRL 488
Dd	425 DRKTAFIGSLNFPDRSAAALTEMGILLFECAPIITRTDLLFAEETGENSYHLRDNRNRI 484
Qy	489 QWHDPATRKTYP--NEPEAKLWKRIAAILSLPIEGLL 525
Dd	485 YWDFIENEKYGVNNYESPFNRFAFIKIISWLPESQL 523

RESULT 32

Q98NZ3	RHILO	
ID	Q98NZ3	RHILO PRELIMINARY; PRI; 466 AA.
AC	Q98NZ3	
DT	01-OCT-2001	(TrEMBLrel. 18, Created)
DT	01-OCT-2001	(TrEMBLrel. 18, Last sequence update)
DT	01-OCT-2001	(TrEMBLrel. 25, Last annotation update)
DE	Phospholipase D-family protein.	
DE	OrderedLocusNames=mlr19675;	
GN	Rhizobium loti (Mesorhizobium loti).	
OS	Rhizobium loti (Mesorhizobium loti).	
OG	Plasmid pMlb.	
OC	Bacteria; Proteobacteria; Alphaproteobacteria	
OC	Phyllobacteriaceae; Mesorhizobium.	
OX	NCBI_TaxID=381;	
FN	[1]	
RP	NUCLEOTIDE SEQUENCE.	
RC	STRAIN=MAFF303099;	
RX	MEDLINE=21082930; PubMed=11214968;	
RA	Kaneko T., Nakamura Y., Sato S., Asamizu E.,	
RA	Watanabe A., Idegawa K., Ishitawa A., Kawashiri	
RA	Kishida Y., Kiyokawa C., Kohara M., Matsumoto	
RA	Takeuchi Y., Nakayama S., Nakazaki N., Shimizu	
RA	Kakeuchi C., Yamada M., Tabata S.	
RT	"Complete genome structure of the nitrogen-fixing Mesorhizobium loti";	
RL	DNA RES. 7:331-338(2000).	
DR	EMBL; AP003017; BAB54862.1; -; Genomic DNA.	
DR	GO; GO:0003824; F:catalytic activity; IEA.	

01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Cardiolipin synthetase family protein.
GN OrderedLocusNames=mlr8077;
OS Rhizobium loti (Mesorhizobium loti).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Phyllobacteriaceae; Mesorhizobium.
OX NCBI_TaxID=381;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=MAFF303099;
RX MEDLINE=21082930; PubMed=11214968;
RA Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
RA Watanabe A., Ideawa K., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,
RA Mochizuki Y., Nakayama S., Nakazaki N., Shimo S., Sugimoto M.,
RA Takeuchi C., Yamada M., Tabata S.;
RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium
RT Mesorhizobium loti";
RL DNA Res. 7:331-338(2000).
DR EMBL; BA000012; BAB53717.1; -; Genomic DNA.
DR GO; GO:0003824; F:catalytic activity; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR001736; PLD.
DR Pfam; PF00614; PLDC; 2.
DR SMART; SM00155; PLDC; 2.
DR PROSITE; PS00035; PLD; 2.
KW Complete proteome.
SQ SEQUENCE 466 AA; 52117 MW; 693301E508A0A788 CRC64;

Query Match 29.6%; Score 804.5; DB 2; Length 466;
Best Local Similarity 41.0%; Pred. No. 2.6e-48;
Matches 193; Conservative 61; Mismatches 180; Indels 37; Gaps 10;

QY 72 GLSDIYLLDDPHEALAAALIESAEHSLDLQYIWRNDISGRLLFNLMYLAERGVVR 131
DB 16 GKSGLLISNDYDAFAARVAARCAARTLDMYLVHDDHGRLLLEQVVAARQGVVR 75
QY 132 LLLDDNNTRGLDLLLLALDSDHPNIEVRLFNFPVLNRK---WRALGYLTFDFPRLNRRMHKS 188
DB 76 MLDDDVNPKSDAAYLALSNNHPNIELKLFNPSGIRARGLRGAELVLLRLPALTRWHKA 135
QY 189 FTADNRATILGRRNIGDEYFKVGEDTFVADLILATGSSVGEVSHDFDRIYASHAHNAT 248
DB 136 WIADDNTAIYVGRNVDGADFAAE-TNFRDLMLLGFVQQAQIFQTFEWCODA---- 190
QY 249 RIIRSGNIGKGLQALGYNDETSRHALLRYRETVESQSPLYKIQ-----TGRIDW 297
DB 191 -----KPIAELGAAAGSHAPYFEGREEKTESTLLSGIRDKGSIABFISASSNVHW 241
QY 298 QSVQTRLISDTPAKGLDRDRKPPVIAGRLODALKQPEKSVYLVSPYFVPTKSGTDALAKL 357
DB 242 VE-RVRVISDPPEK-VRGWRPSRLMKELLPVQSARKVEIVSPYFIPGKSGKSLGDL 299
QY 358 VQDGIDVTVLNSLQATDVAHVSGYVYKRPKLLKAGIKLYELQPNHAVPATKDKGLTGS 417
DB 300 VDDGVQAVLTNSLAATDVAHVGAANYRKRLLRMGVQLFELQPFSSRQPKI---SVFGS 356
QY 418 SVTSLHAKTFIVDGKRIFGSNLDPESARLNTMGVVIESP-K-IAQEMERTLADTTPY 476
DB 357 KGASLHTKASVDRNIGFVGSFNDPKRSVLSNEMGVLFEDENLVAELRRHFKEIAPEA 416
QY 477 AYRVTLDKHNRLOWH--DPATRTKY PNEPEAKLWKRIAALKILSLPIEGLL 525
DB 417 SYRLEL-KNEVLRWHSDEGLQTYTTEPEAAWFRRLAALVRLPIESQL 466

RESULT 35

Q6FB53 ACIAD PRELIMINARY; PRT; 520 AA.
AC Q6FB53;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)

Putative phospholipase D protein.
DE OrderedLocusNames=ACIAD1881;
GN Acinetobacter sp. (strain ADP1).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Moraxellaceae; Acinetobacter.
OX NCBI_TaxID=62977;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=ADP1;
RX PubMed=15514110; DOI=10.1093/nar/gkh910;
RA Barbe V., Vallenet D., Fonknechten N., Kreimeyer A., Oztas S.,
RA Labarre L., Cruveiller S., Robert C., Duprat S., Wincker P.,
RA Ornston L.N., Weissenbach J., Marliere P., Cohen G.N., Medigue C.;
RT "Unique features revealed by the genome sequence of Acinetobacter sp.
RT ADP1, a versatile and naturally transformation competent bacterium.";
RL Nucleic Acids Res. 32:5766-5779(2004).
DR EMBL; CR543861; CAG68710.1; -; Genomic DNA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR001736; PLD.
DR Pfam; PF00614; PLDC; 2.
DR SMART; SM00155; PLDC; 2.
DR PROSITE; PS00035; PLD; 2.
KW Complete proteome.
SQ SEQUENCE 520 AA; 60385 MW; 1DD312966EA749A0 CRC64;

Query Match 29.2%; Score 794; DB 2; Length 520;
Best Local Similarity 37.7%; Pred. No. 1.7e-47;
Matches 186; Conservative 76; Mismatches 177; Indels 54; Gaps 11;

QY 66 HTPHN--NGLSDIYLLDDPHEALAAALIESAEHSLDLQYIWRNDISGRLLFNLMYLA 123
DB 49 HTDENLQKGTAVIPLDDAFISASRIYLIRHAKHQIDLQYIWHNDFVQQLILNELKA 108
QY 124 AERGVRVRLLDNNTRGLDLLLLALDSDHPNIEVRLFNFPVLKRWALGYLTFPRLNRR 183
DB 109 ADRGVKIRLLIDQNGTRLDSTFQALLQHPLEIKVFENPYKFRKLFVDFLFRPNQINHR 168
QY 184 MNKSTADNRATILGRRNIGDEYFKVGEDTFVADLILATGSSVGEVSHDFDRIYASHS 243
DB 169 MNKLLIADNTIATVGTGRNISSEYFDASVQFQTFDLDILFYTAASQAEKVQTFWDSPL 228
QY 244 AHNATRIIRSGNIGKGLQALGYNDETSRHALLRYRETVQ--SPLYKIQITGRID---- 296
DB 229 SVTTQKLITKGPNDLLQLRS-----KYKE-IERVDTPTEKDINEAQLEALNEAL 276
QY 297 -----WQSVQTRLISDTPAKGLDRDRKPPVIAGRLODALKQPEKSVYLVSPYFVPTKS 349
DB 277 KQRYSVKW--AKAHFVADSPKKIHGAIGHAIGNELIYNQMFSGIMHPQKHLELVSAFYVPTQA 334
QY 350 GTDALAKLVQDGDIVTLNSLQATDVAHVSGYVYKRPKLLKAGIKLYELQ----- 402
DB 335 GTQYLSNLQQDVKIRVLTNSFAANDVAIVHATYSQTRKELIKNGIQLYEFKPIIDRKEP 394
QY 403 -----NHAVPATKDKGLTGSVTSLSHAKTFIVDGKRIFGSNLDPESARLNTMGV 455
DB 395 TWYKMTGRVIPA-----KGKSSSLHAKFFDVGCK-VFIGSFNDFPSAYLNTVEGLV 447
QY 456 IESPKIAEQMERTLADTTPYAYRVTLDKHNRLOWHD-PATRTK--YPNPEAKLWKRIA 512
DB 448 VESEDLQNETTKVLDEYLPYIAVQLKLDKNGNLVWLDHNDGQTIEYHHPDPTTRFORM 507
QY 513 AKTLLSLPIEGLL 525
DB 508 MKAVSYLPYIEWMM 520

RESULT 36

Q9HTP4_PSEAE PRELIMINARY; PRT; 529 AA.
AC Q9HTP4;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)

```

DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein.
GN OrderedLocusNames=PA5310;
OS Pseudomonas aeruginosa.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=287;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=ATCC 15692 / PA01;
RX PubMed=20437337; PubMed=10984043; DOI=10.1038/35023079;
RA Stover C.K., Pham X.-O.T., Erwin A.L., Mizoguchi S.D., Warrener P.,
RA Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
RA Garber R.L., Goltry L., Tolentino E., Westbrook-Wadman S., Yuan Y.,
RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
RA Reizer J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.;
RT "Complete genome sequence of Pseudomonas aeruginosa PA01, an
RT opportunistic pathogen.";
RL Nature 406:959-964 (2000).
DR EMBL; AE004943; AAG08695.1; -; Genomic_DNA.
DR F02983; F82983.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR GO; GO:0008654; P:phospholipid biosynthesis; IEA.
DR InterPro; IPR001736; PLD.
DR Pfam; PF00614; PLDC; 2.
DR SMART; SM00155; PLDC; 2.
DR PROSITE; PS00035; PLD; 2.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 529 AA; 59139 MW; DABF5ED5881E63DE CRC64;

Query Match 29.1%; Score 792; DB 2; Length 529;
Best Local Similarity 36.7%; Pred. No. 2.4e-47;
Matches 200; Conservative 86; Mismatches 189; Indels 70; Gaps 16;

QY 21 RSLISLLCLLCSSWLPPL-LEERTSRHNTSKPVLLDNILQIRHTP 68
DB 15 RFLLLALLSGCASTPFPSSALPAEGTWLAKQAE-----IQGRDHP 59

QY 69 HNNGLSDIYLDHPHEALAAALIESAHSLSLDQYIWRNDISGRLLFNLMYLAAGRV 128
DB 60 ---GSGFHLSSASEDAFVARAALIRAQRSLDIQYIVHDLITRALAYELLKAADRGV 116

QY 129 RVRLLDNNTRGLDLDLSDHNPNIIVRNFNPFVRK----WRALGYLTDTPKLNRRM 184
DB 117 RVRIIDDTASDGMWYEIGVLSAHPNIQVRLFNPLHLGRATGITRGVGRLENLSQOHRM 176

QY 185 HNKSFADNRATILGGRNIGDEYFKVGEDTVFADLDILATGVSVEVSHDFDRYASHA 244
DB 177 HNKMLADGTAIVGRLNGDEYFNAPKPEMFTDLDLGVGPANQLSHSFDQYWNLSAIS 236

QY 245 HNAT----RIIRSGNIGKGLQALGYNDTSRHALLRY--RETVEQS----PLYQKIOTGR 294
DB 237 RPIDFLWRAPYPGEL-----ESARKLQYLRKESKESGYIRHLFDGQDQR 285

QY 295 I-DWQS----VQTRLISDTPAKGLDRDRKPP--IAGRLQALQKPEKSVYLVSPFYPT 347
DB 286 LGNMLENTWARAEAIWDAPLKVLSRGEPDPHLLLSPLHAGLFKGQKELILVSAFYVPA 345

QY 348 KSGTDLAKLVQDGDIVTVLNTSLQATDVAAVHSGYVKYRPLKAGIKLYELQPNHVP 407
DB 346 KDLGNLYLTKADSGVRVRLNLSLEATDVPVHAGYAPYRMALLEHGVKLYELRANPOQ 405

QY 408 AT-KDKGLTGSVTSLSHAKRTFVQDKRIFIGSFNLDPRSRNLNTEMGVVISPKIAEQME 466
DB 406 LSGAPWRLHGSSASLSHKAMVDFRKFVIGSFNFDPSILWNTEGVIVDSPLLAQVR 465

QY 467 R-TLADTTPYAYRVTLQKHN----RLQWHD--PATRKTYPNPEPAKLMWKRIAATLSLLP 520
DB 466 QLALEGMAFVSQYVRIDRSRRLVWIDRDGQAQVLRHEP--GSLWRRRLNNAWVAGMIG 524

QY 521 IEGLL 525

Db 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein.
GN OrderedLocusNames=PA5310;
OS Pseudomonas aeruginosa.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=287;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=ATCC 15692 / PA01;
RX PubMed=20437337; PubMed=10984043; DOI=10.1038/35023079;
RA Stover C.K., Pham X.-O.T., Erwin A.L., Mizoguchi S.D., Warrener P.,
RA Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
RA Garber R.L., Goltry L., Tolentino E., Westbrook-Wadman S., Yuan Y.,
RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
RA Reizer J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.;
RT "Complete genome sequence of Pseudomonas aeruginosa PA01, an
RT opportunistic pathogen.";
RL Nature 406:959-964 (2000).
DR EMBL; AE004943; AAG08695.1; -; Genomic_DNA.
DR F02983; F82983.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR GO; GO:0008654; P:phospholipid biosynthesis; IEA.
DR InterPro; IPR001736; PLD.
DR Pfam; PF00614; PLDC; 2.
DR SMART; SM00155; PLDC; 2.
DR PROSITE; PS00035; PLD; 2.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 529 AA; 59139 MW; DABF5ED5881E63DE CRC64;

Query Match 28.5%; Score 774; DB 2; Length 550;
Best Local Similarity 36.1%; Pred. No. 4.8e-46;
Matches 195; Conservative 81; Mismatches 216; Indels 48; Gaps 13;

QY 21 RSLISLLCLLCSSWLPPL-ERTESRHNTSKPVLLDNILQIRHTPHNNGLSDIYLL 79
DB 24 RALVLCALALSGCATPPATTLERTVSHALPPDASTPLADALAVQARAH-PGESGFVVL 82

QY 80 DDPHEALAAALIESAHSLSLDQYIWRNDISGRLLFNLMYLAAGRVVRLLDNNNT 139
DB 83 PRGDEALQMRITAVARAATKLDIQYIAAEDTTGKLLGAALYAADRGVVRMLVDALNF 142

QY 140 RGLDLDLALSDHNPNIIVRNFNPFVRKRW----RALGYLTDTPRLNRRMHNKSFADNRA 195
DB 143 KDIDKLMALDAHANLEVRNFENPFAGPLGNFARTANVFTRIDNFTREHMKAMISDNQI 202

QY 196 TILGGRNIGDEYFKVGEDTVFADLDILATGVSVEVSHDFDRYASHAHNATRIIRSGN 255
DB 203 AIVGGRNLGDEYFNASPTLQFRDLVLAAGVPVTRAVSASFDAYWSSALTYPLPALNRRY 262

QY 256 IGKGLQALGYNDTSRHALLRY-----RETVEQSPLYQKIOTGRID--WQSVQREL 304
DB 263 DAKDLDA-----ARDALRAHWRANATFYNAKPLNATPLAAQIARNEGLVWASAE--F 313

QY 305 ISDTPAK-GLDRDRRKPPIAGRLQALQKPEKSVYLVSPFYPTKSGTDLAKLVQDGDID 363
DB 314 TADSPKLTAAAPDDSYKSPMQQLFALTDQAQREFLVLSFYVPHDAGYNALGRLTARGVR 373

QY 364 VTVLNTSLQATDVAAVHSGYVKYRPLKAGIKLYELQPNHVPATKDKGLTGS--SVTSL 422
DB 374 VAILTNSLAATDAITAVQAQYAPYRVPMLERGVELYEYKPD---PGRSRIGMLGRSRASL 430
```

```
Qy 423 HAKTIVDGRKIFIGSNLDRSARLNTMGVWVIESPKIAEQMERTADTT--PEYAYRVT 481
Db 431 HAKAYVIDRKILVIGSMNLLDRSARLNTMGVWVIESPKIAEQMERTADTT--PEYAYRVT 490
Qy 482 LDKHN-----RLQWHDPAIR--KTPNPEPEAKLWKRIAAKILSLPIEGLL 525
Db 491 LAPDTPGAQTTGAGAPAWPLVWTEIADGQVRYTSDPNAGFYRNLLTGLCLLPIDDQL 550

RESULT 38
Q88CA5_PSEPK
ID Q88CA5_PSEPK PRELIMINARY; PRT; 517 AA.
AC Q88CA5;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DE 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Phospholipase D family protein.
GN OrderedLocusNames=PP5276;
OS Pseudomonas putida (strain KT2440).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=160488;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22423060; PubMed=12534463;
RX DOI=10.1046/j.1462-2920.2002.00366.x;
RA Nelson K.E., Weinel C., Paulsen I.T., Dodson R.J., Hilbert H.,
RA Martins dos Santos V.A.P., Fouts D.E., Gill S.R., Pop M., Holmes M.,
RA Brinkac L.M., Beanan M.J., DeBoy R.T., Daugherty S.C., Kolonay J.F.,
RA Madupu R., Nelson W.C., White O., Peterson J.D., Khouri H.M.,
RA Hance I., Chris Lee P., Holtzapfle E.K., Scanlan D., Tran K.,
RA Moazzez A., Utterback T.R., Rizzo M., Lee K., Kosack D., Moestl D.,
RA Wedler H., Lauber J., Stjepandic D., Hohelsel J., Straetz M., Heim S.,
RA Kiewitz C., Eisen J.A., Timmis K.N., Duesterhoeft A., Tuemmler B.,
RA Fraser C.M.;
RA "Complete genome sequence and comparative analysis of the
RT metabolically versatile Pseudomonas putida KT2440."
RL Environ. Microbiol. 4:799-808(2002).
DR EMBL; AE016793; AAM70841.1; -; Genomic_DNA.
DR TIGR; PP5276; -;
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR GO; GO:0008654; P:phospholipid biosynthesis; IEA.
DR InterPro; IPR001736; PLD.
DR Pfam; PF00614; PLDC; 2.
DR SMART; SM00155; PLDC; 2.
DR PROSITE; PS00035; PLD; 2.
KW Complete proteome.
SQ SEQUENCE 517 AA; 57574 MW; 107E07900AA35751 CRC64;

Query Match 28.4%; Score 773; DB 2; Length 517;
Best Local Similarity 37.8%; Pred. No. 5.2e-46;
Matches 199; Conservative 83; Mismatches 210; Indels 34; Gaps 15;

Qy 21 RSLISLCLLL--CSCSWLPPLEERTESRHFTSKPVLDDNLIQIRHTPHNGLSDIYL 78
Db 5 RAUPLLLVLLGVAGCTSIAP---RETSQALPAHESAFGRSVL--RQAPYGRSGFRL 59
Qy 79 LDDPHEALAAARALIESAEHSLDQYIWRNDISGRLLFNLMYLAERGVVRVRLDDNN 138
Db 60 LPNSNEAFRAAEILRNAQASIDLIQYIVHDGLSTRALVHELLRAADRGVRVRLDDTT 119
Qy 139 TRGLDDLLLDLSDHNPVLEVPFLRK---WRALGYLTDPRNRMHNSFTADNR 194
Db 120 SDGLDVTMGTLDAHPNTHIRVFNPLHLGRSTGVTRAVGRFLNLSRQHRHMKLFLVDNS 179
Qy 195 ATILGGRNIGDEYFKVGEDTVFADLDILATGSGVVGESHDFDRYWASHSAHNATRIIRS 254
Db 180 MAIVGGRNIGDEYDAEPNLFNTDIDLLGVPAEQLGCHSFQVWNSALSRPTDFLWHD 239
Qy 255 NIGKGLQALGVNDSTSHALLRVRETVEQSPLYQKIQT----GRID-WQS-----VQTRLI 305
Db 240 PDANDLRASQRLEVS---LAKAR--TQRKALYDRLMAYQSPRLDWRNELIWAHAQAL 294
```

RESULT 39

```
Q4LJD9_9BURK
ID Q4LJD9_9BURK PRELIMINARY; PRT; 540 AA.
AC Q4LJD9;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DE 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Phospholipase D/Transphosphatidyase precursor.
GN ORFNames=Bcen2424DRAFT_1235;
OS Burkholderia cenocepacia HI2424.
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Burkholderiaceae; Burkholderia; Burkholderia cepacia complex.
OX NCBI_TaxID=331272;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX STRAIN=HI2424;
RG US DOE Joint Genome Institute (JGI-PGF);
RA Copeland A., Lucas S., Lapidus A., Barry K., Dettler C., Glavina T.,
RA Hammon N., Israni S., Pitluck S., Richardson P.;
RT "Sequencing of the draft genome assembly of Burkholderia cenocepacia
RT HI2424."
RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RX STRAIN=HI2424;
RG US DOE Joint Genome Institute (JGI-ORNL);
RA Larimer F., Land M.;
RT "Annotation of the draft genome assembly of Burkholderia cenocepacia
RT HI2424."
RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
CC -! CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AAHL01000070; EAM16198.1; -; Genomic_DNA.
KW Signal.
FT SIGNAL.
SQ SEQUENCE 540 AA; 59105 MW; EA4326D0970D659D CRC64;
```

Query Match 28.4%; Score 772.5; DB 2; Length 540;
Best Local Similarity 36.8%; Pred. No. 6e-46;
Matches 198; Conservative 83; Mismatches 208; Indels 49; Gaps 18;

```
Qy 19 KTRSLISLCLLCSGSLPPLE-ERTESRHFTSKPVLDDNLIQIRHTPHNGLSDIY 77
Db 21 QARLVVCALLPLAACATHTPATSLDRPVSHALSATATPLRALAAPEAAH-PGQSGFR 79
Qy 78 LDDPHEALAAARALIESAEHSLDQYIWRNDISGRLLFNLMYLAERGVVRVRLDDNN 137
Db 80 LLADGATALQMRALARAATKTLDMQYIATEDTTCKLLGAALYAADRGVRVRLVDDL 139
Qy 138 NTRGLDDLLLDLSDHNPVLEVPFLRK----WRALGYLTDPRNRMHNSFTADN 193
Db 140 NFHDIDRVMAALNTHQNIIEIRVFNPFEGASQGRMMERTANFFTRIDSFTTRMHKMIADN 199
Qy 194 RATILGGRNIGDEYFKVGEDTVFADLDILATGSGVVGESHDFDRYWASHSAHNATRIIRS 253
```

